DOCUMENT RESUME

ED 137 380 95 TM 006 186

AUTHOR Marliave, Richard; And Others
TITLE Alternative Procedures for Collecting Instructional Time Data: When Can You Ask the Teacher and When Must You Observe for Yourself?
INSTITUTION Far West Lab. for Educational Research and Development, San Francisco, Calif.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE [Apr 77]
NOTE 120p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977) ; Hard copy not available due to print quality of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Academic Achievement; Attention Span; Classroom Environment; *Classroom Observation Techniques; Comparative Analysis; Correlation; *Data Collection; *Educational Researchers; Elementary Education; Elementary School Students; *Elementary School Teachers; Grade 2; *Instruction; Mathematics Instruction; Predictive Validity; Reading Instruction; Records (Forms); Reliability; Teacher Behavior; *Time
IDENTIFIERS Beginning Teacher Evaluation Study

ABSTRACT

Records kept by teachers and data collected by trained observers were examined as alternatives for measuring allocated and engaged instructional time in elementary school classes. The teacher ratings of student engagement did not correlate with observer ratings of engagement, but a high correlation between teacher ratings of engagement and student aptitude was found. However, teacher records of allocated instructional time were judged to be acceptable, using observed data as a criterion. In general, the predictive validity of observational records is higher than that for teacher records, while that of engaged time is higher than that for allocated time. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

Session # 23.17

ED137380

ALTERNATIVE PROCEDURES FOR COLLECTING INSTRUCTIONAL TIME DATA: WHEN CAN YOU ASK THE TEACHER AND WHEN MUST YOU OBSERVE FOR YOURSELF?

Richard Marliave
Charles W. Fisher
Nikola N. Filby

Far West Laboratory for Educational Research
and Development
1855 Folsom Street
San Francisco, California 94103

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Richard Marliave

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM006 186

Paper presented at the meetings of the American Educational Research Association, New York City, New York, April 4-8, 1977.

The development of this paper was supported, in part, by the Beginning Teacher Evaluation Study, conducted for the California Commission for Teacher Preparation and Licensing, under funds provided by the National Institute of Education.

ABSTRACT

This paper compares teachers and trained observers as alternative sources of instructional time data. These sources are compared in terms of both allocated and engaged instructional time, where allocated time is that instructional time designated or assigned by the teacher and engaged time is that instructional time during which the student is actually attending to the task.

Data were collected in six second grade classes. Teacher records of allocated time were obtained over an eight-week period, while direct observation was conducted for a two-week period. In addition, seven days of paired observation were carried out by the two observers of this study. The results of the paired observation showed inter-observer reliability for engaged instructional time in reading and mathematics to be at levels suitable for most research purposes.

Analyses showed that teacher records of allocated instructional time were positively correlated at a reasonably high level with both allocated and engaged time obtained by direct observation. Adjustments of the teacher records using teacher estimates of student engagement rates did not generally increase the correlations between teacher and observer sources of data. Further analyses of teacher estimates of student engagement rates showed these estimates to be more highly correlated with a general measure of student achievement in reading than they were with estimates of student engagement rates obtained by observation.

Analyses of the relationship between instructional time and student learning led to the tentative conclusion that teacher and observer sources of instructional time data are comparable in terms of their association with student learning. These analyses also indicated that observed estimates of student engagement rates and combinations of instructional time categories could both be used to increase the association between teacher records of allocated time and student learning.

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	METHODOLOGY AND INSTRUMENTATION	5
III.	RESULTS AND DISCUSSION	35
IV.	SUMMARY AND CONCLUSIONS	68
	REFERENCES	72
	APPENDIX A: Reading Content Categories and Examples of Teacher Logs	73
	APPENDIX B: Mathematics Content Categories and Examples of Teacher Logs	84
	APPENDIX C: Teacher Rating Forms	92

LIST OF TABLES

		Page
2.1	Reading: Interobserver correlations for content category by setting combinations.	26
2.2	Mathematics: Interobserver correlations for content category by setting combinations.	27
2.3	Estimates of average class engagement during reading instruction for eight Grade 2 classes.	30
2.4	Summary of data collected on nine Grade 2 classes.	32
3.1	Reading: Means, standard deviations, and correlations for Class 1 from teacher logs and observation.	37
3.2	Reading: Means, standard deviations, and correlations for Class 2 from teacher logs and observation.	38
3.3	Reading: Means, standard deviations, and correlations for Class 3 from teacher logs and observation.	39
3.4	Reading: Means, standard deviations, and correlations for Class 4 from teacher logs and observation.	40
3.5	Reading: Means, standard deviations, and correlations for Class 5 from teacher logs and observation.	41
3.6	Reading: Means, standard deviations, and correlations for Class 6 from teacher logs and observation.	42
3.7	Mathematics: Means, standard deviations, correlations, and average differences for Class 1 from teacher logs and observation. . . .	48
3.8	Mathematics: Means, standard deviations, correlations, and average differences for Class 2 from teacher logs and observation. . . .	49
3.9	Mathematics: Means, standard deviations, correlations, and average differences for Class 3 from teacher logs and observation. . . .	50
3.10	Mathematics: Means, standard deviations, correlations, and average differences for Class 4 from teacher logs and observation. . . .	51
3.11	Mathematics: Means, standard deviations, correlations, and average differences for Class 5 from teacher logs and observation. . . .	52
3.12	Mathematics: Means, standard deviations, correlations, and average differences for Class 6 from teacher logs and observation. . . .	53
3.13	Reading: Means, standard deviations, and intercorrelations for estimates of student engagement in six Grade 2 classes.	59
3.14	Mathematics: Means, standard deviations, and intercorrelations for estimates of student engagement for six Grade 2 classes. . .	62

LIST OF FIGURES

		Page
2.1	Teacher Log Format	12
2.2	Attendance/Group Composition Record	13
2.3	Observation Coding Form	20

I. INTRODUCTION

There is currently a great deal of interest among educational researchers in the relationship between instructional time and student learning. Many of the recent studies of instructional time have been field-based studies that attempt to measure the time devoted to various areas of instruction in ongoing classrooms. However, different studies have used radically different methods for collecting data on instructional time. School records of attendance and class schedules were used by Wiley and Harnischfeger (1974). Records kept by teachers were used by Gump (1967). Various observational procedures with a time base were used by Flanders (1970) and by Stallings and Kaskowitz (1974). Despite this broad range of data collection methodologies for the study of instructional time, little is known about the relative merits of these different data sources. The present paper attempts to augment existing knowledge of alternative procedures for collecting instructional time data by comparing records kept by teachers and data collected by trained observers.

The comparison of teachers and observers as sources of instructional time data is of considerable practical and theoretical importance to researchers. With a moderate amount of training, teachers are able to maintain records of instructional time at little or no cost to the researcher. However, it is impossible to obtain useful inter-teacher reliability data in most ongoing classrooms because there is generally only one teacher involved in a particular unit of instruction. The collection of inter-teacher reliability data outside of the ongoing classroom (in a laboratory setting) would be difficult and costly.

Furthermore, reliability determined in an artificial setting might not generalize to records kept by teachers in their own ongoing classrooms.

Data collected by trained observers is generally subject to more direct control by the researcher than is data obtained from classroom teachers. Inter-observer reliability can be readily obtained in the on-going classroom setting. Also, different observers can be assigned to the same classroom so that possible sources of bias are averaged with each other. In addition, the researcher can usually conduct more extensive training and monitoring of observers under his/her direct employ than would be possible with regular classroom teachers. The outside observer is also more likely to perceive instruction objectively than is the teacher who is directly involved in that instruction. However, trained observers are an expensive commodity. As the period of training and data collection is increased, the cost is also increased. Outside observers also have the disadvantage of being obtrusive to the very phenomena that they must observe. The more detailed the observational records of student activities, the more the observer must move around the room and examine student activities closely.

An additional complication of the measurement of instructional time is the distinction between "allocated" and "engaged" instructional time. Allocated time refers to that amount of time provided or set aside by the teacher for some instructional activity. Engaged time refers to that portion of the allocated time during which the student actually attends to the instructional activity. Presumably, learning can only occur when the student is attending to or actively engaged in the instructional task. Allocation without engagement would not be related to learning outcomes.

However, engaged time is usually more difficult to obtain than is allocated time. First, it is sometimes difficult to determine whether or not a student is actively engaged in an ongoing task. Second, complete data on engaged time requires the continuous observation of individual students. Otherwise, estimated rates of engagement must be determined over shorter instructional periods and applied to longer periods. Considering the importance of measuring engaged instructional time and the complexities of doing so, a major purpose of this paper is the comparison of teacher and observer data on student engagement.

The data to be examined for this paper involve teacher records and direct observation of both allocated and engaged instructional time. This yields four basic measures: teacher records of allocated time, teacher records of engaged time, observation data on allocated time, and observation data on engaged time. In addition, examination is made of data combining teacher and observer sources, wherein teacher records of allocated time have been adjusted according to observationally measured rates of student engagement.

The data to be reported were collected as part of the Beginning Teacher Evaluation Study (BTES) for the 1975-1976 academic year. This paper examines data collected during the fall semester of 1975 for reading and mathematics instruction in second grade classrooms. Teacher records of allocated time were obtained during the fall term over an eight-week period between the administration of a pretest and posttest. Direct observation of both allocated and engaged time was conducted over a two-week period within the same term, with additional pre- and posttesting for this two-week period. Teacher ratings of the average

daily proportion of allocated time during which a student was actually engaged were obtained for each student individually (once during the term). In addition, observational ratings of the proportion of allocated time during which students were actively engaged were obtained by two procedures: 1) the ratio of engaged time to allocated time was calculated using the two-week observational data referred to above, and 2) the proportion of allocated time during which students were engaged was coded directly by a time-sampling observation procedure conducted on one day during the term in each class.

The procedures for obtaining teacher records and for collecting observational data of instructional time are described in detail in the second section of this paper ("Methodology and Instrumentation"). This section also includes data on the inter-observer reliability of the observational procedures. The third section ("Results and Discussion") examines the relationship between instructional time as measured using teacher records and as measured by direct observation. Teacher and observer estimates of student engagement rates are compared. Finally, the third section includes a brief discussion of teacher records and observer data in terms of their relative validity for predicting student learning outcomes.

II METHODOLOGY AND INSTRUMENTATION

Sample

The field work carried out by Far West Laboratory during the continuation year of Phase III-A of the BTES (see Far West Laboratory, 1975) involved a sample of 33 teachers. This sample was composed of 16 Grade 5 and 17 Grade 2 teachers. Each volunteered to participate in the one-year study. It was decided to conduct the reading and mathematics studies with separate samples of teachers. The teachers at both grade levels chose to participate in either the reading or the mathematics sample.

The study reported here concentrated on the Grade 2 reading subsample, which consisted of nine teachers. Given practical and financial constraints, it was not possible to carry out extensive direct observation in all classes. As a result, six of the Grade 2 reading classes were selected for direct observation. Selection for the observation subsample was made on the basis of variety of instructional organization across classes and representation of inner city, suburban, and mixed populations. All of the teachers selected agreed to be included in the observation subsample. Since the direct observation required observers to be present for the entire school day, it was feasible to collect information on both reading and mathematics instruction. As a result, this subsample of six classes was treated as a regular part of the Grade 2 reading sample of nine classes; but, in addition, several mathematics scales were administered to the classes, and teachers kept logs of both reading and mathematics instruction. Therefore, the data for this subsample includes both observer and teacher records of instructional time.

Design

The primary purpose of the data collection conducted during Phase III-A of the BTES was to describe naturally occurring variations in allocated and engaged instructional time, and to relate these variations to growth in student achievement. The strategy was to assess student achievement in a number of content areas on two occasions; once early in the fall and once late in the fall. In the intertest interval, records of allocated time were kept. The intertest period was chosen in such a way that a maximum interval was available without inconvenience to schools during the first two weeks of classes or the week preceding Christmas vacation. It was also necessary to have approximately ten days at the beginning of the school year for contacting teachers and instructing them in procedures for keeping records of allocated time. These practical time constraints determined that the first testing occasion (referred to as occasion A) take place during the first week of October, 1975. Records of allocated time were kept for eight weeks of instruction, after which the second testing (occasion B) was conducted during the first week of December, 1975.

In addition to the records of allocated time, data were collected on engaged time by direct observation. The data on engaged time served two main purposes; first, it allowed estimation of the proportion of allocated time during which students were actively engaged; and second, it provided data for relating student engaged time to achievement. For the second of these purposes, it seemed particularly important to assess engaged time over several successive days, rather than a sample of days. In this way, the engaged time in a particular subject area could be assessed relatively

accurately. Each class was therefore observed for two weeks. In an attempt to create optimal conditions for the assessment of the relation between engaged time and achievement, additional achievement tests were administered at the beginning of the first observation day and at the end of the last observation day. These testing occasions are referred to as OA and OB respectively. The procedure provided 100 percent coverage by direct observation of in-school instruction for every student during the OA-OB period. Observation was carried out by two observers; therefore only two classes could be observed during any one two-week period. As a result, classes were observed in pairs during successive two-week periods within the A-B period.

In summary, all classes[1] were tested during the first week of October, 1975 (occasion A); allocated time records were kept for eight weeks; and then all classes were tested again during the first week of December, 1975 (occasion B). Engaged time for each class was assessed by direct observation during a two-week period, with associated pretests and posttests (occasions OA and OB). The timing of the observation periods was staggered in such a way that pairs of classes were observed during the same two-week period. All classes were observed between testing occasions A and B.

Instructional Process Variables

The instructional process data consisted of measures of both allocated and engaged time spent in particular reading and mathematics content areas. Within content areas, several instructional settings were distinguished.

[1] The data set described here is a subset of the data collected during the continuation of Phase III-A of the Beginning Teacher Evaluation Study (Far West Laboratory, 1975).

Data were collected for every student in each of the participating classes. Allocated time was assessed by a log-keeping procedure and engaged time was assessed by direct observation. The present section of this paper describes the subject-matter and setting categories, the teacher log procedure, and the direct observation procedure. The final portion of this section describes procedures used in deriving two alternative indices of student engagement.

Subject-matter and instructional setting categories. Since instruction is planned and implemented by content area, and since student achievement is most often differentiated by content area, instructional time was first partitioned by content category. Subareas of reading (e.g. decoding, word meaning, comprehending main ideas) and mathematics (e.g., addition with regrouping, subtraction with regrouping, place value) constitute the categories. Reading and mathematics content categories were developed at two levels; general and specific. They were derived from a logical analysis of Grade 2 reading and mathematics objectives, textbooks, and curriculum materials. The original categories were modified and refined by classroom teachers during piloting.

For Grade 2 reading, 10 general content categories were defined. These break down into sixty-eight specific content categories.[2] (All reading content categories are listed in Appendix A.) For grade 2 mathematics, 10 general content categories were defined. These break down into twenty-seven specific content categories.[2] (All

[2]The category systems had a primary use related to the study of test reactivity (Filby & Dishaw, 1976). For this purpose the categories were designed to encompass the entire Grade 2 reading and mathematics curricula.

mathematics content categories are listed in Appendix B.) Specific content categories were developed so that allocated time could be recorded in relatively narrow categories. However, it was not possible to use all of these categories in direct observation. As a result, the general content categories were also devised. In some cases, a general content category corresponds to one specific content category; in most, several specific categories make up one general category.

Within the content categories, broad instructional settings were defined by three fundamental instructional characteristics: adult involvement, pacing, and group size.

The teacher-involvement facet had two elements. Settings in which students worked directly with a teacher (or other adult) were distinguished from settings in which a teacher's primary attention was not directed toward the students being considered. This facet is important because the impact of a teacher's interactive behaviors and skills operates in the former but not the latter type of setting. (The term "teacher" was used in the broad sense, to include any adult directly involved in instruction.) If a class was divided into two groups at some point in time, and one of the groups was engaged in an addition drill with the teacher while the other group was doing seatwork, the students in the drill activity were in a setting with direct teacher involvement. The students who were doing seatwork were in a setting which did not involve a teacher directly, even though the teacher may have occasionally addressed one or more of them. If students were engaged in seatwork, and the teacher's main activity consisted of going from student to student to check or explain work, the teacher was characterized as directly involved, even though he did not interact with all students in the group.

The pacing facet was included to distinguish between settings in which students proceeded at their own pace and settings in which they worked at a pace determined by the teacher (or some other characteristic of instruction). Pacing is very much a matter of degree; students never completely determine their own pace, nor is pace totally determined by external factors. Nevertheless, instructional settings vary considerably in this respect; and, as a result, the rate of student learning may be strongly affected. As a crude operationalization of pacing, a distinction was made between seatwork and group work. Seatwork is the most frequently occurring setting in which students have relatively high control over pace; group work is the situation which is most externally paced.

The third facet of instructional setting was group size. The facet has been the subject of much research and has great intuitive appeal. It was included here, not because of its potential direct effect on learning, but because different group sizes provide the opportunity for very different kinds of student activities, teacher behaviors, and group climates. The mere fact that a student is working in a small group does not imply that a particular kind of instruction <u>will</u> occur; it does act as a necessary (but not sufficient) condition for certain highly-valued teacher behaviors. For instance; the smaller the group, the more closely a teacher can approximate a tutoring situation with each student. However, a lecture to a group of five children is probably very much like a lecture to a group of thirty-five children. Group size, like the other facets of setting, was coded as a dichotomy. Large groups were defined to contain ten or more students; small groups, nine or fewer. (Pilot experience showed that a lower value for the upper bound of "small groups" would have provided very little discrimination among actual classroom groups.)

Teacher logs. The teacher logs were developed by the staff of the BTES. The logs served as the primary source for collecting data on allocated time. All teachers maintained records of time allocated to reading and mathematics instruction. These were referred to as "teacher logs." The logs provided information on content covered and settings for reading and mathematics instruction, on a daily basis, for groups of students in each class. The time allocated to each instructional setting was recorded, with one or more content categories associated with that setting. In highly individualized classes, teachers recorded the content covered and settings used for each student during reading mathematics instruction.

The teacher log format is presented in Figure 2.1. Each one-page log covered one week of instruction for a single group of students. The names of the students in a given group were designated on the attendance/group composition sheet (shown in Figure 2.2). Each teacher listed his class roster on the left hand side of the attendance/group composition form. For a given week, the teachers then designated the reading and mathematics instruction groups for each student and the daily attendance. This procedure allowed for different grouping patterns in reading and mathematics. It also allowed for changes in the composition of student groups during the study.

Reading content was recorded according to the list of categories in Appendix A. Mathematics content categories are listed in Appendix B. Teachers referred to the list to find appropriate codes for content categories that best described the instruction. Teachers were also provided with glossaries which contained examples of each of the content categories, and were individually trained in the log-keeping procedure. Practice logs were kept by each teacher for up to two weeks before data

TEACHER _____ GRADE _____ READING _____ MATH _____ GROUP _____ WEEK _____

TIME:

ONDAY	CONTENT	
	Adult / No Adult and Seatwork / Other	
	MATERIAL	
TUESDAY	CONTENT	
	Adult / No Adult and Seatwork / Other	
	MATERIAL	
WEDNESDAY	CONTENT	
	Adult / No Adult and Seatwork / Other	
	MATERIAL	
THURSDAY	CONTENT	
	Adult / No Adult and Seatwork / Other	
	MATERIAL	
FRIDAY	CONTENT	
	Adult / No Adult and Seatwork / Other	
	MATERIAL	

Figure 2.1
Teacher Log Format

Figure 2.2

ATTENDANCE/GROUP COMPOSITION RECORD

READING MATH (circle one) Teacher_____ Grade_____

Week of_____

Student's Name	Group	M	T	W	Th	F
1.						
2.						
3.						
4.						
5.						
6.						
7.						
8.						
9.						
10.						
11.						
12.						
13.						
14.						
15.						
16.						
17.						
18.						
19.						
20.						
21.						
22.						
23.						
24.						
25.						
26.						
27.						
28.						
29.						
30.						
31.						
32.						
33.						
34.						
35.						

collection began. The training and glossary were intended to ensure reliable categorization of content from teacher to teacher. Content was recorded using the specific content categories.

In classroom situations, content tended to change more quickly than setting. For this reason, several content categories were often designated for one instructional setting. The starting and ending time for each setting was recorded, thereby providing a record of the instructional time allocated to the content covered in each setting. If several different categories were recorded for one setting, (and therefore one time period), then the teacher specified the time devoted to each content category whenever possible. Otherwise, the total period of time was divided by the number of content categories, yielding an estimated time allocated to each category.

The defining characteristics of instructional settings (adult involvement, pace, and group size) have been described above. Direct involvement of an adult covered a range of activities from lecturing to monitoring independent seatwork. "Adult" referred to any teacher, student teacher, or aide. The same adult was not classified as directly involved in more than one setting at a time. Therefore, an adult would not be classified as directly involved in monitoring seatwork if that were a secondary function of the adult.

Regarding the pacing facet: "seatwork" referred to any setting where students worked independently. Two or more students working together, or an adult tutoring one student, was classified as a group-work setting.

The group size facet was not recorded by teachers. This categorization was made by coders when the teacher logs were returned to the Laboratory for processing. Group size was ascertained by checking the

number of students in a particular group on the attendance/group composition form.

In addition to the information noted above, teachers provided a brief description of the materials used in each instructional setting: the name of a textbook and the pages covered, worksheets used for seatwork assignments, and the like.

In summary: for a given week, each teacher recorded how students were grouped for reading and mathematics instruction on the attendance/group compositon form. Daily absence records were kept on the same form; and if group composition changed during the week, the changes were also reported. On the teacher log form itself, teachers kept daily records for each student group. For each day, time periods were blocked off by vertical lines (drawn by the teacher). The beginning and ending times for a setting were recorded along the top of the form. For each setting, teachers recorded adult involvement, pacing, materials, and content categories. In this way, varied instructional patterns could be recorded on the same form. (Examples of completed teacher logs and attendance/group composition sheets are included in Appendices A and B.)

Where teachers grouped students for instruction, this procedure worked well. However, where instruction was highly individualized, variations were adopted. This most often required the keeping of records for individual students; or, where teachers operated a number of "activity stations," records could be kept for each station.

Since the log procedures were quite new, relatively little was known before the study about their measurement characteristics. Therefore, in order to obtain independent assessments of allocated time, two additional data sources were used. First, Far West Laboratory coders, who transferred

the raw teacher logs into machine-punchable formats, spent one day in each classroom. During that day, the coders completed a log for the instruction that occurred. This log was then available for comparison with the teacher log for the same day. Since there was only one day of coder log per teacher, these data were treated in a clinical manner. Second, at the end of each day of direct observation, the Far West Laboratory observers completed logs. From this data source, seven to nine days of logs were made available for comparison with each teacher's log. The results of these comparisons are presented later in this paper.

Direct observation. Data collection by direct observation served two major purposes. First, direct observation of instruction over a two week intertest period provided the basis for relating achievement to amount of engaged time. Observation of all school instruction during this interval eliminated the problems arising from sampling of a few instructional occasions from a relatively long intertest interval. The observation system was intended to capture all instruction relevant to reading and mathematics in terms of engaged time in content and setting categories, which could then be related to achievement measures. The second purpose of the observation system was to provide independently collected data to compare with the allocated time data from teacherlogs. However, since observers assessed engaged time and teachers reported allocated time, quantitative comparison of these two sources (for purposes of determining the reliability of teacher logs) was difficult. So, in addition to their daily observation task, observers completed an allocated time log of the day's instruction. These were used for comparison to the teacher logs.

In the development of this observation system, the selection of the

level of specificity with which to describe classroom phenomena was a difficult problem. The usefulness and practicality of a content- or setting-category can vary tremendously depending upon the number of facets involved in its definition. The more specific the categories, the more difficult the coding of process data, especially when data are to be collected on every student in a given classroom. A decision must be made to collect either more specific information for a smaller number of students or less specific information on a larger number of students.

In this case, the decision was made to describe instructional settings at a relatively global level, in terms of three dichotomous facets (adult involvement, pacing, and group size). Within these settings, content was noted in relatively specific categories. (The setting facets and content categories have been described above.)

During July and August, 1975, Far West Laboratory staff observed teachers of Grades 2 and 5 in year-round schools operating in Fairfield and Hayward, California. On this occasion, attempts were made to code content in reading and mathematics in a large number of specific categories. It soon became clear that content changed very quickly when specific categories were used. For example, teachers handed out seatwork dittos which included work on a relatively large number of specific mathematics content categories. Clearly, it was impractical to record the amount of time spent on each specific category by each student.

After trying several alternatives, this problem was resolved by redefining the content categories. Since the observation covered a two week period in the fall of the year, attention was restricted to a few specific content categories which were commonly taught during that portion of the school year. The other content areas were collapsed into one broad category.

The observation categories chosen for reading were:

1. decoding-long vowels,
2. other decoding
3. word structure-compound words,
4. other word structure,
5. context clues, word meaning and comprehension
6. reading practice,
7. areas related to reading.

The relationships among the specific, general, and observation content categories in reading are shown in Appendix A.

The observation categories chosen for mathematics were:

1. addition without regrouping,
2. addition with regrouping,
3. subtraction without regrouping
4. subtraction with regrouping
5. place value and expanded notation,
6. other.

The relationships among the specific, general, and observation content categories in mathematics are shown in Appendix B.

Focusing on a small number of content categories made observation much more practical, but did not solve all problems. Experience during piloting indicated that content still changed more quickly than setting variables (for example, group size or adult involvement). Rather than attempt a perfect fit between content categories and the setting variables, more than one content designation was allowed for any particular combination of setting descriptors. These setting descriptors (adult involvement, pacing, and group size) were identical to those used in the teacher logs. The adult-involvement setting facet was coded by using "A" to represent cases where the teacher was directly involved and "N" for all other cases. Pacing was operationalized as seatwork (coded "S") and everything else (coded "O"). Small groups (coded "L") were defined as having nine or fewer students. Settings with ten or more students working on the same activity were designated as large groups (coded "H").

The basic strategy of the system was to code all instruction in reading and mathematics for each student in a classroom. This was done by tracking the time students engaged in particular settings in terms of teacher involvement, pacing, and group size, and subsequently coding the content covered within each setting. (For each setting one or more content categories were recorded.)

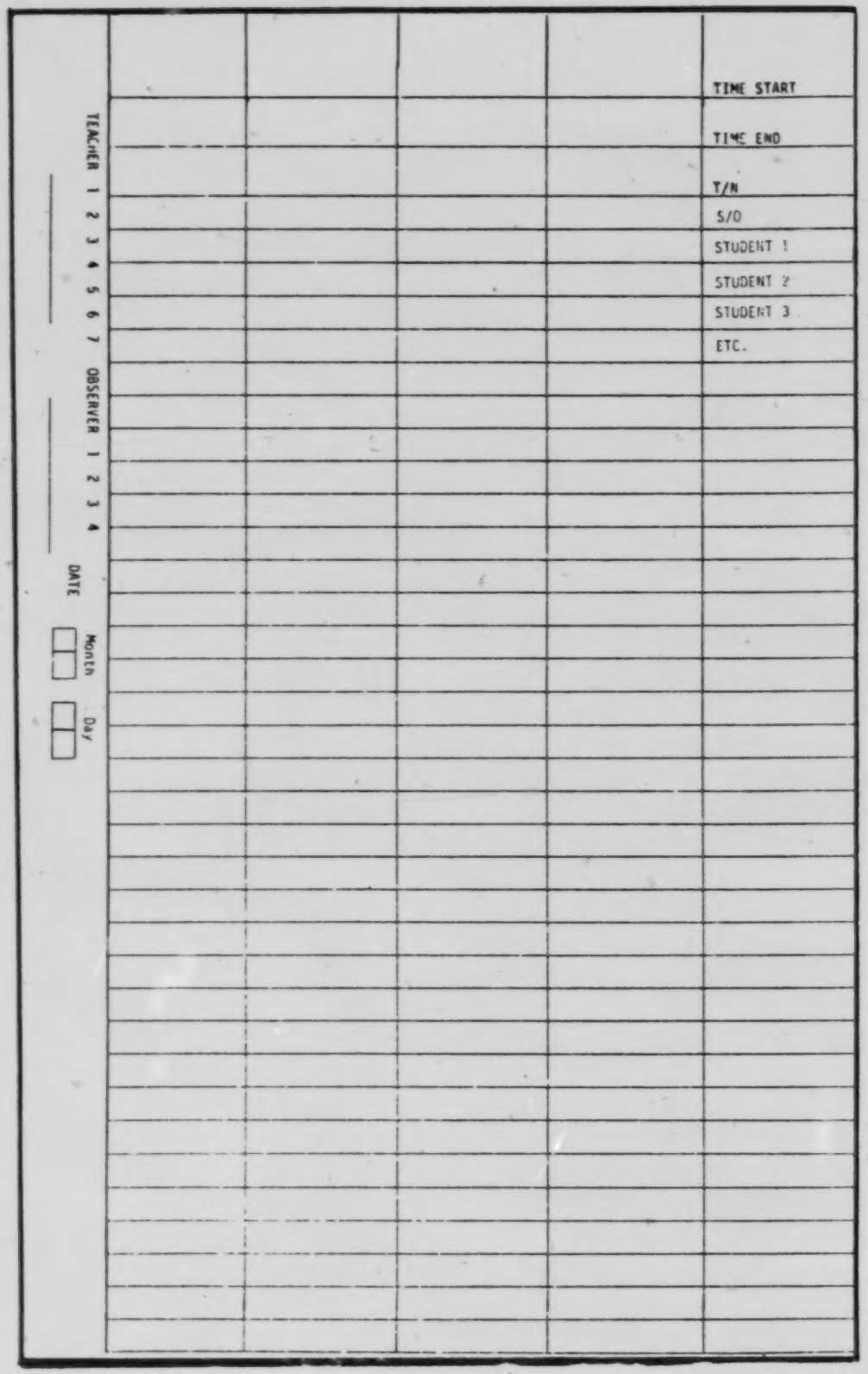
Experience during piloting indicated that one observer could monitor classes of up to thirty students for this information. However, it was essential that the observer know the general routine of the classroom, the materials, and also be able to distinguish one student from another rapidly. These requirements were met by having an observer spend one full day in a class before data collection began. This procedure allowed teacher and students to become accustomed to the observer, and provided the observer with practice in each classroom.

Direct observation procedure. Observers collected data over two consecutive weeks in each classroom. One day was required for memorization of the students' names, and familiarization with the general classroom routine. The remainder of the time (approximately 9 days) was available for official data collection.

Once the observer was familiar with the classroom organization and students, the procedure was relatively straightforward. The observer entered the classroom with the students each morning and used the observation coding form (Figure 2.3) to record data. (The coding form used in the field was 8-1/2 inches by 14 inches. It has been reduced in size for display in Figure 2.3.) Students' names were placed in the columns. The four lefthand columns were used for recording starting and ending times, teacher involvement and pacing codes.

Figure 2.3

Observation Coding Form
(Reduced from 8-1/2" X 14")



The form was used in the following way. The observer noted the starting time for any group setting wherein reading or mathematics was the content. All times were recorded to the nearest minute. Teacher involvement and pacing for each group were then coded in the appropriate columns. Finally, the content was coded in the cell below the name of each student in that setting.

If the content was the same for all students in a setting, then the content was coded for the student appearing first in the list; and a horizontal line was drawn across the appropriate cells for each of the other students in that group. This indicated that the content code was the same for all students in that group. In the simplest case, where a setting came to an end at a particular time for the whole group, the ending time was recorded. If some students in that setting covered different categories of content, then those categories were coded under the names of the appropriate students. If one or more of the students in a setting left that setting, then the end time was entered in the cell for that student directly under the content code. In this way, all students who started out in the same setting could leave it at different times and still be accounted for. If a student entered an existing setting after it started, then the observer coded that student's start time under his name and then coded the content. Thus, if a cell for a particular student began with a time, it was implied that the group time entered in the far left column did not apply to that student. If the last entry in a cell for a particular student was a time, it implied that he left the group before it ended and the end time for the setting (second column from left) did not apply for that student. Similarly, if a student started off in setting A, changed to setting B, and then

returned to setting A again, the sequence could be coded. The cell under the student's name might contain a content code, a time, another time, a content code, and a third time. This configuration would represent a case where the student started the setting with the whole group and was working on the content listed first. This work continued until the first time listed in his cell, at which point the student changed to another setting. At the second time listed in the cell, the student returned to the first setting and worked on the content listed next in the cell. The final time recorded in the cell represents the point at which the student left the setting again, and, in the example being considered, the setting continued to exist after the student's second departure. The time during which the student was not in the setting being discussed could be accounted for by looking in another row on the form (that is, in another setting). However, note that if the interim setting did not involve reading or mathematics, then no entry would have been made for that interval.

To recapitulate: each row on the form represented a setting as defined by teacher involvement and pacing. Several rows could be active at any one time. Content and information which was associated with individual students (as opposed to groups) was recorded in the columns of the form under the names of the particular students. In this way, one observer kept track of all the students in the class. Note that when a student was working on content which was not recordable within one of the categories of reading or mathematics as defined for this study, no codes were recorded.

Group size was not necessarily included in the set of codes. However, the group size for any setting could be recovered from the codes already described. For a particular student at a particular time, group size for

the setting could be determined by examining the row in which the student was included and counting the number of students in the row at the same point in time. (As previously stated: for purposes of analysis, group size was considered a dichotomy; small groups defined as having nine or fewer members, large groups as having more than nine. Since the exact group sizes were available from the raw data, this cutting point could be easily changed for additional analyses.)

The space at the right of the sheet was used for comments or clarifications as they were required. Forms with the names of students were printed for each class. The names of teacher and observer and the date of observation were also recorded on each form.

In carrying out the observation routine, it was necessary for the observer to move about the room to look at materials being worked on by students. Experience showed that the content coding required a thorough knowledge of the materials actually being used by students. This was especially true in cases where the program was highly individualized.

The observation procedure was designed to collect information on engaged time. If students were not engaged in the task at hand, then time was subtracted from each setting for each student depending upon how much time that student was unengaged. When time was subtracted for unengagement, it was done so in multiples of one minute; momentary inattention was ignored.

Engagement was judged by the observer with the aid of several guidelines. When students were working on tasks which required an overt response, engagement was relatively easy to judge. When students were working on tasks which did not involve overt responses, the situation was somewhat more difficult. In the latter cases, observers used student eye

contact and body position as indicators of engagement. If a student was in a discussion group, watching the various speakers in turn and apparently following the discussion, then the time was considered engaged time. If a student was discussing an unrelated topic with other students, or was clearly not attending to the task, then the time was considered unengaged time. The distinction was fairly crude; students were considered unengaged only when the situation was unambiguous.

Observer reliability. The observation data were collected by two observers. After approximately two weeks of training, the observers simultaneously collected data in two classrooms over a four day period for reliability purposes. Both observers went to Class A for two full days and then to Class B for two full days. The data obtained in this period were transferred to the standard coding booklet; and times were collapsed over days, classes and setting codes, so that total engaged times were available for each student for each content category from each of the two observers. Interobserver correlations were computed, and showed good agreement on most content categories.

After this post-training check, the observations were carried out in the study classrooms. This required approximately six weeks. Following the data collection, the observers returned to the same two classes and simultaneously observed Class A for 2 days and Class B for one day.

These data were processed along with those collected at the post-training period. The data were collapsed over the seven days (four pre and three post) of observation, yielding total time in content-by-setting combinations for each student from each observer. Interobserver correlations are discussed below separately for reading and mathematics.

The interobserver correlations for each reading content-by setting combination are presented in Table 2.1. For the calculation of inter-observer agreement indices, the students from both classes were pooled, yielding a sample of 45 students. Some setting-by-content combinations were rarely (or never) observed during the seven day period. This resulted in some correlations being calculated on distributions with very little variance. In some cases only one student had a non-zero engaged time. This accounts for many of the low correlations. Where the distributions were all zeros for both observers, two dashes appear in the table. These represent cases of perfect agreement; that is, neither observer recorded any time for any student in that content-by-setting combination. Where there was a reasonable amount of time recorded, the correlations were relatively high, indicating that engaged time in content-by-setting combinations can be reliably recorded by different observers.

The setting information was used for descriptive purposes only. Time in content areas was used both for descriptive purposes and in analyses of time in content with achievement. The bottom row of Table 2.1 presents the interobserver agreement when the data were collapsed over setting. Note again that the coefficients were relatively high.

The interobserver correlations for each mathematics content-by-setting combination are presented in Table 2.2. The discussion above of Table 2.1 for interobserver correlations in reading applies equally to Table 2.2 for mathematics. As for the reading data, the bottom row of Table 2.2 presents the interobserver agreement when the mathematics data were collapsed over setting. Note again that the coefficients were

Table 2.1

Reading

Interobserver correlations for content category by setting combinations. Data were collected in two classrooms over a total of seven school days. Four of the days occurred after training but before the study data were collected, while three of the days occurred after the study data were collected.

	OBSERVATION CONTENT CATEGORIES						
Setting Combinations	Long Vowels (RL)	Other Decoding (RD)	Compound Words (RC)	Other Word Structure (RS)	Combined Comprehension (RM)	Reading Practice (RP)	Areas Related To Reading (RO)
ASH	.55	.45	.40	.30	.85	.89	.82
ASL	-.04[a]	.43	.00[a]	-.05[a]	-.04	.41	-.06[a]
AOH	1.00	.93	--	--	.93	1.00	1.00
AOL	.99	.62	--	.00[a]	.23	.69	1.00
NSH	.00[a]	--	--	.00	.00	.92	--
NSL	--[b]	.16[a]	--	--	.08[a]	.22	-.06[a]
NOH	--	--	--	--	--	--	--
NOL	--	1.00	--	--	--	-.12	--
All Settings Combined	.95	.91	.63	.94	.85	.64	.97

Note Number of subjects = 45

A = adult directly involved
S = seatwork
L = low group size
N = no adult directly involved
O = other (non-seatwork)
H = high group size

[a] These coefficients represent cases where only a few students had non-zero times assessed by one or both observers. Seven of the coefficients had between five and eight students with non-zero times, while the others had three students with non-zero times.

[b] a -- indicates perfect agreement between observers but all students had zero recorded time.

Table 2.2

Mathematics

Interobserver correlations for content category by setting combinations. Data were collected in two classrooms over a total of seven school days. Four of the days occurred after training but before the study data were collected, while three of the days occurred after the study data were collected.

	Content Categories					
Settings	Addition Without Reg.	Addition With Reg.	Sub. Without Reg.	Sub. With Reg.	Place Value	Other Matn.
ASH	.75	.00*	.70	.00*	.59	.78
ASL	.00*	--	.00*	--	.00*	.00*
AOH	.00	--	.00	.93	--	1.00
AOL	.88	--	.79	--	--	.98
NSH	--[a]	--	--	--	--	--
NSL	.78	--	--	--	--	.00*
NOH	--	--	--	--	--	--
NOL	.77	--	.00*	--	--	--
All settings combined	.71	.00*	.87	.47	.69	.93

Note: Number of subjects = 45

A = adult involved
S = seatwork
L = low group size
N = no adult involved
O = other (non seatwork)
H = high group size

[a] A -- indicates perfect agreement between observers but all students had zero recorded time.

* These coefficients represent cases where only a few students had non-zero times assessed by one or both observers. One of the starred coefficients has 5 students with non-zero times, while the others have 3 students or fewer with non-zero times.

relatively high. (The coefficient for subtraction with regrouping (0.47) was an exception. Very little time was recorded in this content category and the low correlation reflects the lack of variance.)

Student engagement rates. Although the direct observation procedure provided information on the amount of engaged time students spent in a two week instruction period, there was no direct information available on student engagement rates. Two methods of estimating engagement rates were tried.

As noted earlier, observers completed a log at the end of each day of observation. This log contained the amount of allocated time in reading and mathematics for students in the class for a particular day. For most classes, there were seven full days of instruction for which both allocated time from observer logs and engaged time from direct observation were available (one class had six days). For each student, the total time allocated to reading and mathematics and the total engaged time in reading and mathematics were calculated (over the 6 or 7 day period). An observed engagement rate was then computed for reading and for mathematics for each student by taking the ratio of total engaged time in reading to total time allocated to reading and likewise for mathematics.

Since the observed engagement rate could be computed only after extensive observation of each student, it was desirable to find an alternative procedure that would be less expensive. The alternative procedure was based on adjusted teacher ratings of student attentiveness. Teachers were asked to rate each student in terms of the percent of the time that the student paid attention during class. These ratings were made twice:

once for instructional settings where an adult was directly involved, and once for settings where no adult was directly involved. The percent attentiveness ratings were made by placing a check in one of nine categories, where each category represented an increment of 10 percent on a 0 percent to 100 percent scale. (The directions to teachers and the two rating forms are included in Appendix C. By an oversight, the category representing 31 to 40 percent was omitted from the form.)

The teacher ratings of attentiveness were assigned the mid-category value; that is, a check in the 81-90 percent category was assigned a value of 0.85. This provided a distribution of attentiveness scores for each class. However, comparison from one class to another would be hazardous, since errors due to teachers' tendencies to rate high or low would appear as between-class differences. In an attempt to correct for possible teacher bias, class estimates of mean engagement were made.

The estimates were based on data collected during instruction in reading. An observer visited each class for one day. During the reading instruction periods, the observer counted the number of students engaged and the total number of students nominally working on reading. This procedure was repeated every four minutes. In this manner, average class engagement estimates were calculated. The reliability with which students were coded for engagement in different classes was acceptable, ranging from .65 to .85. The results of this procedure are shown in Table 2.3.

These average class engagement estimates were used to adjust the teacher ratings of student engagement. The adjustment was made in such a way that each adjusted class mean was equal to the average class engagement estimate. The adjustment is specified in the following equation:

Table 2.3

Estimates of average class engagement during reading instruction for eight[a] Grade 2 classes.

Class	Average Number of Students Observed	Number of Time Samples	Time Sample Interval (Mins.)	Average Engagement[b]
1	15	18	4	.44
2	11	44	4	.49
3	19	27	4	.25
4	7	41	4	.59
5	15	23	4	.41
6	16	31	4	.51
7	13	33	4	.55
8	7	38	5	.50

[a] Although there are nine classes in the sample, this procedure was carried out in classes 1 through 8. No data are available for class number 9.

[b] These estimates were calculated from one day of observation per class. In all cases data were collected during class time which was allocated to reading activities. Since teachers allocate varying amounts of time to reading, the time period covered by the observation differs considerably. The observers counted the number of students engaged at four minute intervals (with one exception). They recorded the number of students engaged, the time, and the number of students in the classroom who were part of the BTES study and who were nominally working on reading activities. The average engagement was calculated by summing the number of students engaged over the total number of time samples and dividing by the sum of the number of students in the classroom being followed by BTES and nominally working on reading activities. No distinctions have been made between setting combinations or subareas of content within reading.

$$Y_{ij} = \frac{\overline{E_j}}{\overline{R_j}} R_{ij}$$

where Y_{ij} is the adjusted teacher rating of attentiveness for student i in class j, R_{ij} is the teacher rating of student attentiveness, $\overline{R}_j$ is the class mean of the teacher ratings of student attentiveness for class j, and $\overline{E}_j$ is the mean class engagement estimate for class j. This procedure prevents Y_{ij} from being negative, and preserves the relative ranking of students within class.

Data Collection

The data collected for the nine Grade 2 classes are summarized schematically in Table 2.4. Note that the classes numbered one through six comprise the observation subsample for which data are presented in this paper. The eight-week test data were comprised of the scores obtained from testing occasions A (first week of October) and B (first week of December). The teacher log data describe the reading and mathematics instruction for the A-B intertest period. The two-week test data were comprised of scores obtained on testing occasions OA and OB. For Classes 1 and 3, this period fell in the latter half of October. For the remaining two pairs of classes, (numbers 4 and 5 and numbers 2 and 6), the OA-OB interval came during the first and last two weeks of November respectively. The direct observation data and observer log data describe reading and mathematics instruction during the OA-OB interval.

Table 2.4

Summary of data collected on nine Grade 2 classes.

Class	2 week observation data	2 week test scores	8 week log data	8 week test scores	8 week attitude data	2 week observer logs	1 day coder logs	1 day coder estimate of mean class engagement	teacher ratings of student attentiveness
1	R,M	R,M	R,M	R,M	R,M	R,M	R	R	R
2	R,M	R,M	R,M	R,M	R,M	R,M	R	R	R
3	R,M	R,M	R,M	R,M	R,M	R,M	R	R	R
4	R,M	R,M	R,M	R,M	R,M	R,M	R	R	R
5	R,M	R,M	R,M	R,M	R,M	R,M	R	R	R
6	R,M	R,M	R,M	R,M	R,M	R,M	R	R	R
7			R	R	R,M		R	R	R
8			R	R	R,M		R	R	R
9			R	R	R,M		--[a]	--[a]	R

Notes

R represents reading data; M represents mathematics data.

Classes 1 through 6 comprise the observation subsample for which data are presented in this paper.

[a] The one day coder log and the coder estimate of mean class engagement were not obtained for class 9.

At occasions A and B, the reading battery was administered in four 45-minute group testing sessions and the mathematics tests were administered in a group situation over one 30-minute period. They were administered by Far West Laboratory staff, but not by the observers. At the OA and OB occasions, both mathematics and reading tests were administered in one 45-minute session. This testing was administered by the observers. All test administrators were briefed on the testing procedure, and approximately half of the testers administered at least one of the tests in a classroom practice session before testing began. Those test administrators who did not have a practice administration acted as observers at least once while a test was being administered to a class. Test administrators completed testing report forms after every administration, and were debriefed after testing occasions A and B.

Data collection for the teacher logs began early in September with one-to-one meetings with each participant. Materials on log-keeping were explained, and teachers began to keep practice logs up to two weeks before the A testing occasion. Each teacher was visited several times so that any questions about log-keeping could be answered. The amount of feedback which teachers required varied considerably. The more complex the organization for instruction was, the more complicated the log-keeping became.

Once the class rosters were finalized and teachers had some practice, the log-keeping seemed to go smoothly. Teachers were asked to complete their logs each day, and to return them to the Far West Laboratory by mail every Friday. This procedure worked quite well, although teachers were sometimes late in returning logs, and it is not certain that all teachers completed them at the end of each day.

Data collection for the direct observation procedure be[illegible] with

arrangements with the teacher for a two week period for observation. At this time, the observer obtained an outline of classroom routine and discussed the nature of the observation with the teacher. It was made clear to the teacher that information on engaged time in instruction for individual students would be collected, and that no data on teacher behavior were being recorded. The OA and OB testing was discussed, and teachers were told what scales would be administered. In addition, teachers were asked to spend time on instruction in decoding long vowels and in place value. This request was intended to ensure that all students would have at least some time in a common content category. It was desirable to have significant amounts of engaged time in one or more time categories; otherwise it would be difficult to demonstrate growth in achievement over a two-week period.

On the first observation day, the observer memorized the names of students who were to be observed, and became familiar with the classroom routine. During this day, the observation procedure was practiced in this new setting, and teacher and students had time to become accustomed to the observer. Every day during the observation period, the observer entered the class with or before the students and remained for the entire school day. This allowed the coding of all instruction relevant to reading and mathematics. On the second day of the observation period, the OA testing was administered by the observer. Immediately after the testing, observation data collection began and continued during in-school hours until the OB testing date.

III RESULTS AND DISCUSSION

Characteristics of Teacher Allocated Time Logs

The teacher logs provided measures of allocated time over the A-B period. At a practical level, the procedure proved workable. Teachers were able to use the content and setting categories, and to keep records of time allocated to various kinds of instruction. The procedure was also flexible enough to allow data collection in very different classroom organizational structures. Comparison of the teacher logs with observer logs provided information on the accuracy of the recorded allocated times. The observer logs had been completed at the end of each school day during the OA-OB period. (This task was a secondary priority for the observers, since all of their in-school time was taken up with direct observation; and after school hours, their primary task was the transference of direct observation data from the observation coding form to the standard coding booklets.)

The allocated time logs completed by the observers differed from the teacher logs in at least two important ways. First, observer allocated time logs recorded content at the level of general content categories, while the teachers' logs used specific content categories. This mismatch prevented the comparison of allocated time within all of the specific content categories, but did allow comparison of allocated time within the general content categories.

Second, the observer logs were coded for content using a strategy referred to as "focus coding." This requires that an instructional activity be placed in one particular content category, if possible -- the most complex category which describes the activity. (Multiplication

would be coded as "multiplication," and not as part multiplication and part addition, even though addition is part of the multiplication process.) When contents were covered in sequence, each was coded with its appropriate allocated time, but where contents were coextensive in time, the more complex content was coded. This focusing on one content category was used in direct observation and hence carried over into the coding of the observer allocated time logs.

The teacher logs, on the other hand, used a strategy referred to as "multiple coding" for categorizing content. In this procedure, teachers are encouraged to use more than one content code, if it improves the description of the activity. In processing the logs, if an instructional activity received more than one content code for a time interval, the time was distributed equally over the content codes. The same activity can be coded quite differently, depending upon whether focus coding or multiple coding is used. In spite of these differences, the observer logs were the best source of information for checking the accuracy of teacher allocated time logs. Data are presented below comparing teacher records of instructional time with observer records. These data are presented and discussed first for reading, then for mathematics.

Tables 3.1 through 3.6 present comparative data on observer and teacher logs of reading instruction. Each table presents information on one teacher. The tables are identical in format. These tables deal only with the content information of the logs.

The tables are based on teacher logs, observer logs, and direct observation information for days when all three sources were available for a given class. Table 3.6, describing the log characteristics for

Table 3.1

Reading

Means, standard deviations, and correlations for Class 1 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 16)

Class 1

Content Category	A Allocated Time from Teacher Logs	B Adjusted Allocated Time from Teacher Logs	C Allocated Time from Observer Logs	D Engaged Time from Direct Observation	r_{AC}	r_{AD}	r_{BD}
Long vowels (GCC 1)[a]	30 (15)	13 (7)	40 (27)	24 (23)	.63	.85	.87
Total decoding (GCC 1,2)	262 (75)	122 (63)	256 (65)	176 (40)	-.77	.42	.32
Compound words (GCC 4)	58 (31)	24 (14)	2 (3)	19 (35)	.24	.85	.67
Total word structure (GCC 4,5)	126 (43)	55 (23)	47 (24)	59 (41)	.30	.96	.69
Reading practice (GCC 9)	121 (14)	53 (16)	431 (57)	71 (20)	-.48	.00	.19
Other reading (GCC 8)	6 (11)	2 (4)	69 (35)	27 (12)	-.96	-.80	-.72
Context clues (GCC 3)	56 (20)		140 (58)		.89		
Word meaning (GCC 6)	33 (6)		13 (8)		-.84		
Comprehension of text (GCC 7)	26 (16)		107 (23)		-.50		
Total comprehension (GCC 3,6,7)	114 (7)	50 (17)	260 (81)	76 (43)	-.06	-.07	.39
Total reading (GCC 1 through 9)	629 (53)	282 (104)	1062 (148)	410 (85)	.67	.70	.42

[a] General content category numbers are shown in parentheses.

Table 3.2

Reading

Means, standard deviations, and correlations for Class 2 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 18)

Class 2

Content Category	A Allocated Time from Teacher Logs	B Adjusted Allocated Time from Teacher Logs	C Allocated Time from Observer Logs	D Engaged Time from Direct Observation	r_{AC}	r_{AD}	r_{BD}
Long vowels (GCC 1)[a]	8 (11)	4 (6)	13 (16)	4 (5)	.98	.82	.83
Total decoding (GCC 1,2)	254 (52)	124 (28)	188 (48)	104 (35)	.97	.74	.73
Compound words (GCC 4)	1 (1)	0 (1)	2 (3)	3 (5)	1.00	.91	.91
Total word structure (GCC 4,5)	17 (19)	9 (9)	11 (11)	14 (15)	.98	.91	.92
Reading practice (GCC 9)	118 (33)	58 (18)	88 (37)	132 (44)	.56	.69	.56
Other reading (GCC 8)	85 (21)	41 (11)	44 (16)	25 (12)	.69	.51	.47
Context clues (GCC 3)	7 (9)		0 (0)		.00		
Word meaning (GCC 6)	0 (0)		0 (0)		--[b]		
Comprehension of text (GCC 7)	81 (26)		17 (9)		.30		
Total comprehension (GCC 3,6,7)	88 (23)	42 (10)	17 (9)	21 (11)	.57	.29	.31
Total reading (GCC 1 through 9)	562 (110)	273 (60)	348 (86)	296 (89)	.95	.84	.76

[a] General content category numbers are shown in parentheses.

[b] A -- indicates perfect agreement between sources of time information however there was no variance on either variable.

Table 3.3

Reading

Means, standard deviations, and correlations for Class 3 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 20)

Class 3

Content Category	A Allocated Time from Teacher Logs	B Adjusted Allocated Time from Teacher Logs	C Allocated Time from Observer Logs	D Engaged Time from Direct Observation	r_{AC}	r_{AD}	r_{BD}
Long vowels (GCC 1)[a]	77 (10)	19 (6)	100 (9)	38 (4)	-.35	.30	.55
Total decoding (GCC 1,2)	223 (21)	55 (15)	295 (52)	106 (18)	.53	.41	.29
Compound words (GCC 4)	38 (18)	9 (4)	24 (6)	11 (4)	.55	.23	.47
Total word structure (GCC 4,5)	39 (18)	9 (4)	24 (6)	12 (4)	.57	.38	.59
Reading practice (GCC 9)	267 (36)	66 (19)	175 (30)	49 (12)	.29	.16	.57
Other reading (GCC 8)	46 (41)	12 (12)	5 (7)	2 (0)	-.75	.00	.00
Context clues (GCC 3)	0 (0)		0 (0)		--[b]		
Word meaning (GCC 6)	0 (0)		0 (0)		--		
Comprehension of text (GCC 7)	120 (62)		38 (54)		.98		
Total comprehension (GCC 3,6,7)	120 (62)	31 (20)	38 (54)	16 (20)	.98	.96	.93
Total reading (GCC 1 through 9)	695 (79)	172 (51)	537 (60)	184 (27)	.94	.65	.81

[a] General content category numbers are shown in parentheses.

[b] A -- indicates perfect agreement between sources of time information however there was no variance on either variable.

Table 3.4

Reading

Means, standard deviations, and correlations for Class 5 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 26)

Class 5

Content Category	A Allocated Time from Teacher Logs	B Adjusted Allocated Time from Teacher Logs	C Allocated Time from Observer Logs	D Engaged Time from Direct Observation	r_{AC}	r_{AD}	r_{BD}
Long vowels (GCC 1)[a]	10 (9)	5 (5)	0 (0)	3 (5)	.00	-.16	-.11
Total decoding (GCC 1,2)	73 (29)	31 (15)	83 (34)	45 (30)	.68	.41	.12
Compound words (GCC 4)	0 (0)	0 (0)	6 (6)	4 (11)	.00	.00	.00
Total word structure (GCC 4,5)	16 (26)	7 (11)	26 (9)	20 (15)	.34	.09	.11
Reading practice (GCC 9)	114 (50)	45 (18)	141 (48)	55 (25)	.99	.76	.47
Other reading (GCC 8)	16 (15)	8 (8)	37 (17)	17 (11)	.65	.59	.58
Context clues (GCC 3)	14 (9)		0 (0)		.00		
Word meaning (GCC 6)	14 (15)		0 (0)		.00		
Comprehension of text (GCC 7)	70 (50)		127 (33)		.33		
Total comprehension (GCC 3,6,7)	98 (44)	43 (24)	127 (33)	80 (24)	.55	.19	.22
Total reading (GCC 1 through 9)	318 (64)	132 (48)	414 (64)	217 (46)	.29	.37	.01

[a] General content category numbers are shown in parentheses.

Table 3.5

Reading

Means, standard deviations, and correlations for Class 4 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 14)

Class 4

Content Category	A Allocated Time from Teacher Logs	B Adjusted Allocated Time from Teacher Logs	C Allocated Time from Observer Logs	D Engaged Time from Direct Observation	r_{AC}	r_{AD}	r_{BD}
Long vowels (GCC 1)[a]	20 (17)	13 (12)	27 (20)	6 (10)	.74	.71	.64
Total decoding (GCC 1,2)	90 (33)	54 (25)	186 (53)	195 (52)	.48	.45	.39
Compound words (GCC 4)	23 (7)	14 (5)	9 (1)	13 (3)	.04	.25	.44
Total word structure (GCC 4,5)	23 (7)	14 (5)	16 (3)	13 (3)	.04	.39	.53
Reading practice (GCC 9)	143 (35)	84 (28)	184 (50)	127 (69)	-.18	.13	.35
Other reading (GCC 8)	8 (11)	5 (7)	73 (41)	40 (35)	.67	.49	.55
Context clues (GCC 3)	7 (11)		0 (0)		.00		
Word meaning (GCC 6)	65 (24)		0 (0)		.00		
Comprehension of text (GCC 7)	57 (30)		96 (30)		-.23		
Total comprehension (GCC 3,6,7)	129 (52)	78 (36)	96 (30)	37 (16)	-.06	.24	.23
Total reading (GCC 1 through 9)	394 (90)	235 (79)	554 (104)	412 (79)	.66	.55	.63

[a] General content category numbers are shown in parentheses.

Table 3.6

Reading

Means, standard deviations, and correlations for Class 6 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over six days of instruction for which all three sources of time information were available. (N = 18)

Class 6

Content Category	A Allocated Time from Teacher Logs	B Adjusted Allocated Time from Teacher Logs	C Allocated Time from Observer Logs	D Engaged Time from Direct Observation	r_{AC}	r_{AD}	r_{BD}
Long vowels (GCC 1)[a]	32 (21)	18 (14)	19 (5)	18 (22)	.90	.47	.56
Total decoding (GCC 1,2)	201 (32)	103 (43)	166 (49)	79 (26)	.66	.34	.31
Compound words (GCC 4)	1 (6)	1 (4)	2 (7)	1 (3)	1.00	-.06	-.06
Total word structure (GCC 4,5)	3 (8)	2 (5)	18 (12)	6 (7)	.94	.03	.03
Reading practice (GCC 9)	97 (21)	51 (21)	46 (27)	24 (10)	.30	.22	-.09
Other reading (GCC 8)	49 (18)	26 (13)	8 (8)	7 (7)	.35	.33	.48
Context clues (GCC 3)	3 (7)		0 (0)		.00		
Word meaning (GCC 6)	17 (16)		0 (0)		.00		
Comprehension of text (GCC 7)	5 (6)		186 (58)		.14		
Total comprehension (GCC 3,6,7)	24 (19)	11 (10)	186 (58)	66 (31)	.24	-.20	.05
Total reading (GCC 1 through 9)	373 (65)	192 (77)	423 (79)	182 (42)	.94	.47	.71

[a] General content category numbers are shown in parentheses.

Class 6, is based on data from six days of instruction. The tables for the other five classes are each based on information from seven days of instruction.

The rows of each table are labeled by general content category. Note that rows 2, 4, and 10 represent subtotals for decoding, word structure and comprehension respectively. The last row presents total time in general content categories 1 through 9. The entries in column A are allocated times from the teacher logs. Column C presents allocated time from the observer logs. Information in all general content categories was included for columns A and C, since the logs provided this information. Columns A and C provide the basic comparison for allocated time. Column D presents engaged time from direct observation. Since the direct observation system used one content category to cover general content categories 3, 6 and 7, some rows in column D are blank. Allocated times from the teacher logs were multiplied by the adjusted teacher ratings of student attentiveness (discussed previously). This product, referred to as "adjusted allocated time from teacher logs" is presented in the tables as column B.

The purpose of calculating the adjusted allocated time was to allow comparison with the engaged time from direct observation. Therefore, column B presents data in only those rows (general content categories) for which engaged time from direct observation was available. Columns B and D then allow a comparison of measures of adjusted allocated and engaged time from independent sources. In addition to comparison of means and standard deviations, three sets of Pearson product-moment correlation coefficients were calculated. The first, r_{AC}, describes

the relation between the two sources of allocated time. The second, r_{AD}, describes the relation between allocated time from teacher logs and engaged time from direct observation. Finally, r_{BD}, represents the degree of relationship between adjusted allocated time and engaged time.

A simple summary of Tables 3.1 through 3.6 is difficult but several comparisons do shed some light, for example, the comparison of allocated time from teacher logs and from observer logs. The means in columns A and C for rows 1 through 10 did not agree consistently; for some rows they seemed to agree quite well, for others they did not. No class had agreement in all rows, but there were several content categories where most classes agreed. In the main, these were categories where relatively little time had been allocated. Of the 60 average differences in columns A and C (6 classes by rows 1 through 10), 25 were less than 15 minutes in magnitude and 32 were less than 30 minutes in magnitude. Note that rows 1 and 2, rows 3 and 4 and rows 7, 8, 9 and 10 are not independent. Therefore some of the "disagreements" between the means of columns A and C were counted twice. In any case there were many large average differences.

The differences between columns A and C for class 1 (Table 3.1) appeared to be larger than those for the other classes, rendering the log data from class 1 less useful than that from other classes. Considering all of the classes, there were several examples of miscategorization while in other cases pieces of the reading program have been included by the observer but not by the teacher or vice versa. These comparisons reflect a number of sources of error. One was the use of the different coding strategies for the two data sources. (Teacher logs were coded

using the multiple coding strategy, while the observer logs used focus coding.) Amount of error due to coding strategy differences as compared to other sources of error is unknown. Class 5 (Table 3.5) demonstrates this difficulty. Note that, in this table, the means in columns A and C match quite well, with the exception of general content categories 3, 6 and 7. However, note too that the sum of general content categories 3, 6 and 7 for column A (98 minutes) is in moderate agreement with the corresponding sum for column C (127 minutes). In this case, the observer log (which used focus coding) allocated all of the time in question to comprehension of text, while the teacher log (using multiple coding) distributed the time (over context clues, word meaning and comprehension of text).

The variance within class in time allocated to reading (over 40 days of instruction) was moderate; that is, students in the same class tended to get more or less similar amounts of time allocated to reading. Differences among students within the same class on total time allocated to reading were due in large part to absenteeism.

The content categories function as a partially ipsative set -- the amount of time in any one category was not independent of the time in the other categories. Furthermore, an error in one category tended to cause errors in one or more additional categories.

For total times allocated to reading, relatively large differences between sources of data were found. Note that when allocated time was summed over content categories, coding strategy differences no longer had an effect. With the exception of class 1, the correlations between columns A and C were moderate to high.

The correlation between the two sources of allocated time data accumulated over general content categories provided a reasonable summary description for the classes. Within five of the classes (1, 2, 3, 4, and 6), the students were ranked similarly on amount of allocated time from the teacher logs and from the observer logs. The correlations for these classes were 0.67, 0.95, 0.94, 0.66, and 0.94 respectively. For Class 5 the correlation was 0.29.

Upon examination of the correlations between allocated time from teacher logs and engaged time from direct observation, several interesting findings emerged. First, for Class 5 (which had poor overall agreement between the two sources of allocated time), the value of r_{AD} was 0.37. For the five classes where the two sources of total allocated time agreed reasonably well within class, r_{AD} was approximately equal to r_{AC} for three classes (1, 2 and 4), and was substantially lower than r_{AC} for the others (Classes 3 and 6).

Turning now to the comparison of adjusted allocated time from teacher logs and engaged time from direct observation: the differences in the means in columns B and D were considerably smaller than those described for columns A and C. In this case, 23 of 42 comparisons within general content categories (represented by rows in Tables 3.1 through 3.6) showed means differing by less than 15 minutes. Agreement across classes was good for long vowels, compound words and total word structure. An important question was whether the allocated time from teacher logs or the adjusted allocated time from the teacher logs was more highly correlated with the engaged time from direct observation. A comparison of r_{AD} and r_{BD} shows that in three classes (3, 4 and 6) the

adjusted allocated time from teacher logs was more highly correlated with engaged time from direct observation. For each of these classes the improvement was substantial. However in two of the remaining three classes (1 and 5) the decrease in the relationship brought about by the adjustment procedure was also substantial. So, in this sample, the characteristics of the allocated time from teacher logs were improved in three of the classes but not improved in the other three by the adjustment procedure.

Tables 3.7 through 3.12 present comparative data on observer and teacher logs of mathematics instruction. Each table presents information on one teacher. The tables are identical in format. These tables deal only with the content information of the logs.

The tables are based on teacher logs, observer logs, and direct observation information for days when all three sources were available for a given class. Table 3.12, describing the log characteristics for Class 6, is based on data from six days of instruction. The tables for the other five classes are each based on information from seven days of instruction.

The rows of each table are labeled by general content category. The first 10 rows represent the 10 general content categories, row eleven is a subtotal for general content categories 6 through 10, and row twelve shows total time in mathematics instruction (the sum over general content categories 1 through 10). The entries in column A are allocated times from the teacher logs. Column C presents allocated time from the observer logs. Information in all general content categories was included for columns A and C, since the logs provided this information. Columns A and C provide the basic comparison for allocated time. Column D presents

Table 3.7
Mathematics

Means, standard deviations, correlations and average differences for class 1 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 16)

CLASS 1

	General Content Category	A Allocated Time From Teacher Logs	B Adjusted Allocated Time (From Teacher Logs)	C Allocated Time From Observer Logs	D Engaged Time From Direct Observation	r_{AC}	r_{AD}	r_{BD}	Average Difference[a] for columns A and C	Average Difference[b] for columns B and D
1	Addition No Regrouping	55 (11)	24 (7)	51 (4)	26 (9)	-.22	-.67	-.18	6 (11)	10 (7)
2	Addition With Regrouping	0 (0)	0 (0)	0 (0)	1 (1)	--[c]	.00	.00	0 (0)	1 (1)
3	Subtraction No Regrouping	53 (11)	23 (7)	32 (2)	18 (7)	.33	-.62	-.25	21 (10)	9 (8)
4	Subtraction With Regrouping	0 (0)	0 (0)	0 (0)	0 (0)	--	--	--	0 (0)	0 (0)
5	Place Value	0 (0)	0 (0)	30 (0)	23 (10)	--	.00	.00	30 (0)	23 (10)
6	Number System	44 (2)		9 (3)		-.10			35 (3)	
7	Measurement	0 (0)		0 (0)		--			0 (0)	
8	Geometry	0 (0)		0 (0)		--			0 (0)	
9	Word Problems	41 (2)		21 (3)		.59			21 (3)	
10	Other	0 (0)		84 (17)		.00			84 (17)	
	Categories 6 through 10 Combined	86 (4)	38 (12)	114 (19)	37 (7)	.98	.62	.25	31 (9)	10 (7)
	Total Mathematics Categories 1 Through 10 Combined	194 (23)	84 (26)	227 (23)	104 (16)	.35	-.54	-.25	39 (17)	31 (23)

[a] The entries in this column are average differences without regard for sign between allocated time from teacher logs and allocated time from observer logs.

[b] The entries in this column are average differences without regard for sign between adjusted allocated time (from teacher logs) and engaged time from direct observation.

[c] A -- indicates perfect agreement between sources of time information but there was no variance on either variable.

Table 3.8

Mathematics

Means, standard deviations, correlations and average differences for class 2 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 18)

CLASS 2

General Content Category	A Allocated Time From Teacher Logs	B Adjusted Allocated Time (From Teacher Logs)	C Allocated Time From Observer Logs	D Engaged Time From Direct Observation	r_{AC}	r_{AD}	r_{BD}	Average Difference[a] for columns A and C	Average Difference[b] for columns B and D
1 Addition No Regrouping	51 (12)	25 (6)	0 (0)	6 (8)	.00	.29	.11	51 (12)	20 (8)
2 Addition With Regrouping	0 (0)	0 (0)	0 (0)	0 (0)	--[c]	--	--	0 (0)	0 (0)
3 Subtraction No Regrouping	55 (15)	27 (8)	5 (1)	7 (4)	-.01	.59	.51	51 (15)	20 (6)
4 Subtraction With Regrouping	0 (0)	0 (0)	0 (0)	0 (0)	--	--	--	0 (0)	0 (0)
5 Place Value	101 (24)	49 (12)	209 (47)	176 (45)	.79	.81	.75	109 (32)	127 (37)
6 Number System	42 (9)		38 (17)		.66			7 (12)	
7 Measurement	0 (0)		0 (0)		--			0 (0)	
8 Geometry	0 (0)		0 (0)		--			0 (0)	
9 Word Problems	0 (0)		0 (0)		--			0 (0)	
10 Other	0 (0)		0 (0)		--			0 (0)	
Categories 6 through 10 Combined	42 (9)	21 (5)	38 (17)	0 (0)	.66	.00	.00	7 (12)	21 (5)
Total Mathematics Categories 1 Through 10 Combined	249 (49)	121 (27)	252 (55)	189 (51)	.84	.85	.70	24 (16)	68 (37)

[a] The entries in this column are average differences without regard for sign between allocated time from teacher logs and allocated time from observer logs.

[b] The entries in this column are average differences without regard for sign between adjusted allocated time (from teacher logs) and engaged time from direct observation.

[c] a -- indicates perfect agreement between sources of time information but there was no variance on either variable.

Table 3.9

Mathematics

Means, standard deviations, correlations and average differences for class 3 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 20)

CLASS 3

General Content Category	A Allocated Time From Teacher Logs	B Adjusted Allocated Time (From Teacher Logs)	C Allocated Time From Observer Logs	D Engaged Time From Direct Observation	r_{AC}	r_{AD}	r_{BD}	Average Difference[a] for columns A and C	Average Difference[b] for columns B and D
1 Addition No Regrouping	64 (7)	16 (5)	51 (4)	23 (8)	.88	.14	-.42	13 (4)	10 (8)
2 Addition With Regrouping	0 (0)	0 (0)	0 (0)	0 (0)	--[c]	--	--	0 (0)	0 (0)
3 Subtraction No Regrouping	71 (13)	17 (6)	70 (16)	44 (13)	.86	.65	.51	4 (7)	26 (11)
4 Subtraction With Regrouping	0 (0)	0 (0)	0 (0)	1 (3)	--	.00	.00	0 (0)	1 (3)
5 Place Value	157 (12)	39 (11)	133 (20)	47 (10)	.54	.27	.68	24 (17)	9 (8)
6 Number System	54 (23)		0 (0)		.00			54 (23)	
7 Measurement	0 (0)		0 (0)		--			0 (0)	
8 Geometry	0 (0)		0 (0)		--			0 (0)	
9 Word Problems	29 (7)		19 (4)		1.00			10 (2)	
10 Other	0 (0)		0 (0)		--			0 (0)	
Categories 6 through 10 Combined	83 (23)	21 (9)	19 (4)	5 (7)	.18	.44	.46	64 (23)	16 (8)
Total Mathematics Categories 1 Through 10 Combined	374 (42)	93 (28)	273 (31)	120 (22)	.92	.54	.45	101 (18)	32 (21)

[a] The entries in this column are average differences without regard for sign between allocated time from teacher logs and allocated time from observer logs.

[b] The entries in this column are average differences without regard for sign between adjusted allocated time (from teacher logs) and engaged time from direct observation.

[c] 0 -- indicates perfect agreement between sources of time information but there was no variance on either variable.

Table 3.10

Mathematics

Means, standard deviations, correlations and average differences for class 4 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 14)

CLASS 4

General Content Category	A Allocated Time From Teacher Logs	B Adjusted Allocated Time (From Teacher Logs)	C Allocated Time From Observer Logs	D Engaged Time From Direct Observation	r_{AC}	r_{AD}	r_{BD}	Average Difference[a] for columns A and C	Average Difference[b] for columns B and D
1 Addition No Regrouping	39 (21)	24 (14)	68 (21)	90 (30)	.63	.43	.38	31 (12)	66 (28)
2 Addition With Regrouping	6 (2)	3 (1)	14 (0)	0 (0)	.00	.00	.00	8 (2)	3 (1)
3 Subtraction No Regrouping	15 (10)	9 (6)	52 (26)	1 (0)	.41	.00	.00	38 (21)	8 (6)
4 Subtraction With Regrouping	0 (0)	0 (0)	0 (0)	0 (0)	--[c]	--	--	0 (0)	0 (0)
5 Place Value	19 (12)	12 (8)	16 (9)	18 (12)	-.17	.10	.05	13 (8)	12 (8)
6 Number System	40 (23)		53 (33)		.11			31 (24)	
7 Measurement	0 (0)		0 (0)		--			0 (0)	
8 Geometry	25 (15)		0 (0)		.00			25 (15)	
9 Word Problems	0 (0)		0 (0)		--			0 (0)	
10 Other	0 (0)		0 (0)		--			0 (0)	
Categories 6 through 10 Combined	65 (34)	40 (22)	53 (33)	1 (6)	.11	.13	.14	37 (25)	39 (22)
Total Mathematics Categories 1 Through 10 Combined	145 (65)	88 (47)	202 (43)	110 (35)	.26	.51	.43	72 (50)	42 (24)

[a] The entries in this column are average differences without regard for sign between allocated time from teacher logs and allocated time from observer logs.

[b] The entries in this column are average differences without regard for sign between adjusted allocated time (from teacher logs) and engaged time from direct observation.

[c] A -- indicates perfect agreement between sources of time information but there was no variance on either variable.

Table 3.11

Mathematics

Means, standard deviations, correlations and average differences for class 5 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over seven days of instruction for which all three sources of time information were available. (N = 26)

CLASS 5

General Content Category	A Allocated Time From Teacher Logs	B Adjusted Allocated Time (From Teacher Logs)	C Allocated Time From Observer Logs	D Engaged Time From Direct Observation	r_{AC}	r_{AD}	r_{BD}	Average Difference[a] for columns A and C	Average Difference[b] for columns B and D
1 Addition No Regrouping	103 (14)	41 (10)	220 (29)	76 (24)	.96	.64	-.01	118 (15)	35 (26)
2 Addition With Regrouping	0 (0)	0 (0)	5 (1)	3 (1)	.00	.00	.00	5 (1)	3 (1)
3 Subtraction No Regrouping	92 (21)	37 (12)	148 (29)	65 (23)	.84	.57	.17	57 (16)	29 (23)
4 Subtraction With Regrouping	0 (0)	0 (0)	0 (0)	0 (0)	--[c]	--	--	0 (0)	0 (0)
5 Place Value	14 (5)	6 (3)	44 (6)	13 (7)	.81	.48	.51	31 (3)	7 (6)
6 Number System	38 (15)		31 (13)		.95			8 (5)	
7 Measurement	0 (0)		0 (0)		--			0 (0)	
8 Geometry	0 (0)		0 (0)		--			0 (0)	
9 Word Problems	8 (7)		0 (0)		.00			8 (7)	
10 Other	0 (0)		0 (0)		--			0 (0)	
Categories 6 through 10 Combined	46 (10)	19 (8)	31 (13)	9 (7)	.93	.83	.76	15 (5)	10 (5)
Total Mathematics Categories 1 Through 10 Combined	254 (37)	104 (31)	448 (59)	166 (43)	.93	.59	.09	195 (28)	63 (49)

[a] The entries in this column are average differences without regard for sign between allocated time from teacher logs and allocated time from observer logs.

[b] The entries in this column are average differences without regard for sign between adjusted allocated time (from teacher logs) and engaged time from direct observation.

[c] a -- indicates perfect agreement between sources of time information but there was no variance on either variable.

Table 3.12
Mathematics

Means, standard deviations, correlations and average differences for class 6 on allocated time from teacher logs, adjusted allocated time from teacher logs, allocated time from observer logs, and engaged time from direct observation. These data are summed over six days of instruction for which all three sources of time information were available. (N = 18)

CLASS 6

General Content Category	A Allocated Time From Teacher Logs	B Adjusted Allocated Time (From Teacher Logs)	C Allocated Time From Observer Logs	D Engaged Time From Direct Observation	r_{AC}	r_{AD}	r_{BD}	Average Difference[a] for columns A and C	Average Difference[b] for columns B and D
1 Addition No Regrouping	22 (22)	9 (9)	78 (26)	44 (21)	.63	.60	.62	56 (20)	35 (17)
2 Addition With Regrouping	4 (5)	2 (3)	0 (0)	0 (0)	.00	.00	.00	4 (5)	2 (3)
3 Subtraction No Regrouping	19 (23)	7 (9)	68 (21)	32 (13)	.56	.53	.53	50 (19)	25 (11)
4 Subtraction With Regrouping	0 (0)	0 (0)	0 (0)	0 (0)	--[c]	--	--	0 (0)	0 (0)
5 Place Value	0 (0)	0 (0)	0 (0)	0 (0)	--	--	--	0 (0)	0 (0)
6 Number System	0 (0)		86 (24)		.00			86 (24)	
7 Measurement	43 (27)		0 (0)		.00			43 (27)	
8 Geometry	22 (29)		0 (0)		.00			22 (29)	
9 Word Problems	30 (28)		0 (0)		.00			30 (28)	
10 Other	0 (0)		0 (0)		--			0 (0)	
Categories 6 through 10 Combined	94 (54)	53 (35)	86 (24)	41 (18)	.85	.92	.80	31 (18)	21 (16)
Total Mathematics Categories 1 Through 10 Combined	139 (23)	72 (28)	232 (52)	117 (33)	.78	.75	.36	92 (37)	45 (35)

[a] The entries in this column are average differences without regard for sign between allocated time from teacher logs and allocated time from observer logs.

[b] The entries in this column are average differences without regard for sign between adjusted allocated time (from teacher logs) and engaged time from direct observation.

[c] A -- indicates perfect agreement between sources of the information but there was no variation on either variable.

engaged time from direct observation. Since the direct observation system used one content category to cover general content categories 6 through 10, some rows in column D are blank. Direct observation data were available for only the first five general content categories and the sum of general content categories 6 through 10. Allocated times from the teacher logs were multiplied by the adjusted teacher ratings of student attentiveness (discussed previously). This product, referred to as "adjusted allocated time (from teacher logs)" is presented in the tables as column B.

The purpose of calculating the adjusted allocated time was to allow comparison with the engaged time from direct observation. Therefore, column B presents data in only those rows (general content categories) for which engaged time from direct observation was available. Columns B and D then allow a comparison of measures of adjusted allocated and engaged time from independent sources. In addition to comparison of means and standard deviations, three sets of Pearson product-moment correlation coefficients were calculated. The first, r_{AC}, describes the relation between the two sources of allocated time. The second, r_{AD}, describes the relation between allocated time from teacher logs and engaged time from direct observation. Finally, r_{BD}, represents the degree of relationship between adjusted allocated time and engaged time. The final columns in the tables provide average differences (without regard for algebraic sign) between columns A and C, and B and D respectively. These average differences give an idea of the discrepancy between the various time measures. No further comment will be made on these last two columns since, in the main, they were almost identical to the difference between the means for their respective columns. This implied that the differences

within a general content category were almost always of the same sign.

A simple summary of Tables 3.7 through 3.12 is difficult but several comparisons do shed some light, for example, the comparison of allocated time from teacher logs and from observer logs. The means in columns A and C for rows 1 through 11 did not agree consistently; for some rows they seemed to agree quite well while for others they did not. No class had agreement in all rows, but there were several content categories where all classes agreed. In the main, these were categories where no time had been allocated (general content categories 2, 4, 7 and 10). Of the 66 average differences in columns A and C (6 classes by rows 1 through 11), 39 were less than 15 minutes in magnitude. However, among the remaining average differences, there were a number of high values. These 66 mean differences were based on comparisons within general content categories (or, in the case of row 11, a subset of general content categories). These comparisons reflect a number of sources of error. One was the use of the different coding strategies for the two data sources. (Teacher logs were coded using the multiple coding strategy, while the observer logs used focus coding.) Amount of error due to coding strategy differences as compared to other sources of error is unknown. Class 2 (Table 3.8) demonstrates this difficulty clearly. Note that, in this table, the means in columns A and C match remarkably well, with the exception of general content categories 1, 3 and 5. However, note too that the sum of general content categories 1, 3 and 5 for column A (207 minutes) is almost identical with the corresponding sum for column C (214 minutes). In this case, the observer log (which used focus coding) allocated practically all of the time in question to place

value, while the teacher log (using multiple coding) divided the time (half going to place value and one quarter each to addition and subtraction without regrouping).

The variance in time allocated to mathematics was relatively low within class; that is, students in the same class tended to get the same amount of time allocated to mathematics. Note that the content categories do function as a partially ipsative set -- the amount of time in any one category was not independent of the time in the other categories. Furthermore, an error in one category tended to cause errors in one or more additional categories.

For total times allocated to mathematics, relatively large differences between sources of data were found. Note that when allocated time was summed over content categories, coding strategy differences no longer had an effect. The correlations between columns A and C were moderately high, where the data allowed the coefficients to be calculated. The major exceptions to this observation occurred in Classes 1 and 4.

The correlation between the two sources of allocated time data accumulated over general content categories provided a reasonable summary description for the classes. Within four of the classes (2, 3, 5, and 6), the students were ranked similarly on amount of allocated time from the teacher logs and from the observer logs. The correlations for these classes were 0.84, 0.92, 0.93, and 0.78 respectively. For Classes 1 and 4, the correlations were 0.35 and 0.26 respectively.

Upon examination of the correlations between allocated time from teacher logs and engaged time from direct observation, several interesting findings emerged. First, for Classes 1 and 4 (which had poor overall agreement between the two sources of allocated time), the values of r_{AD}

were -0.54 and 0.51 respectively. For the four classes where the two sources of total allocated time agreed reasonably well within class, r_{AD} was approximately equal to r_{AC} for two classes (2 and 6), and was substantially lower than r_{AC} for the others (Classes 3 and 5).

Turning now to the comparison of adjusted time from teacher logs and engaged time from direct observation: the differences in the means in columns B and D followed a pattern similar to that described for columns A and C. In this case, 23 of 36 comparisons within general content categories showed means differing by less than 15 minutes. Agreement across classes was good for general categories 2, 4 and 5. An important question was whether the allocated time from teacher logs or the adjusted allocated time from the teacher logs was more highly correlated with the engaged time from direct observation. A comparison of r_{AD} and r_{BD} shows that, in all classes except Class 1 (where both r_{AD} and r_{BD} were negative), the allocated time from teacher logs was more highly correlated with engaged time from direct observation. For Classes 5 and 6, the differences between r_{AD} and r_{BD} were substantial. So, in this sample, the characteristics of the allocated time from teacher logs were not improved by the adjustment procedure.

Summarizing the characteristics of the teacher logs for both reading and mathematics, it appears that the teacher logs of instructional time generally agreed at an acceptable level with the criterion of observational data of instructional time. This was true for both observational data of allocated time and for observational data of engaged time. However, when the teacher logs of allocated time were multiplied by the adjusted teacher ratings of student attentiveness, the resulting products were more often

at a lower level of agreement with observational data of engaged time than were the teacher records alone. Therefore, it may be concluded that the adjusted teacher ratings of student attentiveness were unsuccessful. However, another adjustment procedure, using only observational data, was developed. This procedure is examined below.

Comparison of Engagement Rating Procedures

The ineffectiveness of adjustments of teacher records of allocated instructional time using teacher ratings of student attentiveness suggested that this procedure for rating engagement should be compared with other procedures. As discussed previously, the adjusted teacher ratings of student attentiveness consisted of teacher ratings of student attentiveness adjusted with estimates of class engagement rates obtained by observers on a single day. The adjusted teacher ratings were compared with ratings of student engagement based on the ratio of engaged to allocated time as recorded by observers over the OA-OB period. These comparisons are discussed first for reading data, then for mathematics data.

Table 3.13 presents information on student engagement during reading instruction. Columns A, B, and C, of this table show the means and standard deviations for different engagement indices. The average teacher ratings of student attentiveness (column A) were, with one exception, higher than either of the averages of the indices based on independent observation procedures (columns B and C). The standard deviation (within class) of the teacher ratings (column A) was also greater, with only one exception, than those for columns B and C. With the exception of Classes 2 and 4, there was moderate agreement between columns B and C. The

Table 3.13

Reading

Means, standard deviations and intercorrelations for estimates of student engagement in six Grade 2 classes.

Class	Number of Students	A Teacher ratings of Student Attentiveness[a]	B Adjusted Teacher Ratings of Student Attentiveness[b]	C Observed Engagement Rate[c]	D Academic Status	r_{BC}	r_{BD}	r_{CD}
1	16	.68 (.22)[d]	.44 (.14)	.38 (.05)	160 (91)	.35	.70	.47
2	18	.86 (.08)	.49 (.04)	.85 (.14)	139 (77)	.35	.57	.41
3	20	.58 (.16)	.25 (.07)	.34 (.03)	36 (35)	.79	.42	.28
4	14	.69 (.14)	.59 (.12)	.75 (.10)	147 (95)	-.08	.42	.33
5	26	.68 (.18)	.41 (.11)	.52 (.07)	133 (66)	-.30	.80	-.29
6	18	.62 (.22)	.51 (.18)	.44 (.09)	77 (54)	.39	.58	.42
All students pooled	112	.68 (.19)	.44 (.15)	.54 (.20)	114 (81)	.42	.58	.35

[a] The teacher ratings of student attentiveness are described in section II.

[b] The adjusted teacher ratings of student attentiveness were obtained by multiplying the teacher ratings of student attentiveness by a different constant for each class. The mean of the adjusted ratings equal the mean class engagement determined by one day of observation in each class, as described in section II.

[c] The observed engagement rate was calculated by taking the ratio for each student of engaged time in reading (direct observation) and allocated time in reading (observer logs), as described in section II.

[d] Standard deviations are shown in parentheses.

correlation between the adjusted teacher ratings of student attentiveness and the observed engagement rates varied considerably for the six classes. Since the number of students within classes was small, only one of these correlations (Class 3) appeared to be inordinately large. Thus with the exception of Class 3, the adjusted teacher ratings did not correlate within classes consistently and positively with the observed engagement rates. When all students were pooled, the correlation was positive and moderate in size. Since this correlation coefficient is affected by the fairly large between-class differences, its size is not surprising. Furthermore, the between-class component of the pooled correlation of adjusted teacher ratings with observed engagement rates is attributable to the adjustment alone (which was based on a single day of observation, as discussed previously). The between-class component is not attributable to the teacher ratings themselves, because the mean of the adjusted teacher ratings for a given class is independent of the teacher ratings for that class.

On the other hand, the correlations between the adjusted teacher ratings of student attentiveness and academic status[3] were all positive and large. This could be interpreted in several ways. It may be that aptitude and student attentiveness were strongly related; or more likely, that the teachers' ratings of student attentiveness were strongly biased by teacher perceptions of student aptitude. Note that the within-class correlations in the table were not affected by the adjustment procedure, since the adjustment coefficient was a constant within a given class.

[3]Academic status is a measure of overall student achievement in reading. This may be considered as a measure of student aptitude for school achievement.

Therefore, the within-class correlations for adjusted teacher ratings with academic status are independent of the adjustment factor and attributable entirely to the unadjusted teacher ratings. The observed engagement rates, on the other hand, were derived without reference to student academic status. This table points out that teacher ratings of student attentiveness were more strongly related to academic status than to observed engagement rates. Note that the observed engagement rates were lower in correlation with academic status.

Table 3.14 presents information on student engagement during mathematics instruction. Columns A, B, and C, of this table show the means and standard deviations for different engagement indices. The average teacher ratings of student attentiveness were, in every case, higher than either of the averages of the indices based on independent observation procedures. With the exception of Classes 2 and 3, there was close agreement between columns B and C for both means and standard deviations. (Remember that the observation procedure underlying the adjustment for column B was conducted on reading instruction rather than mathematics instruction; this may make comparisons between the columns hazardous.) The correlation between the adjusted teacher ratings of student attentiveness and the observed engagement rates fluctuated around zero for the six classes. Since the number of students within classes was small, none of these correlations appeared to be inordinately large. Thus, the adjusted teacher ratings did not correlate within classes with the observed engagement rates. When all students were pooled, the correlation was low and positive. As with the analyses for reading, the between-class component of this pooled correlation is independent of the

Table 3.14
Mathematics

Means, standard deviations and intercorrelations for estimates of student engagement for six grade 2 classes.

		A	B	C	D			
Class	Number of Students	Teacher Ratings of Student Attentiveness[a]	Adjusted Teacher Ratings of Student Attentiveness[b]	Observed Engagement Rate[c]	Academic Status	r_{BC}	r_{BD}	r_{CD}
1	16	.[illegible] (.22)[d]	.44 (.14)	.47 (.11)	160 (91)	-.12	.70	-.04
2	18	.86 (.08)	.49 (.04)	.74 (.09)	139 (77)	-.25	.57	-.44
3	20	.58 (.16)	.25 (.07)	.44 (.07)	36 (35)	.15	.42	.05
4	14	.69 (.14)	.59 (.12)	.54 (.13)	147 (95)	.33	.42	.28
5	26	.68 (.18)	.41 (.11)	.37 (.08)	133 (66)	-.18	.80	-.17
6	18	.62 (.22)	.51 (.18)	.51 (.12)	77 (54)	.23	.58	-.06
All Students Pooled	112	.68 (.19)	.45 (.15)	.50 (.15)	114 (81)	.27	.58	.05

[a] The teacher ratings of student attentiveness are described in section II.

[b] The adjusted teacher ratings of student attentiveness were obtained by multiplying the teacher ratings of student attentiveness by a different constant for each class. The mean of the adjusted ratings equals the mean class engagement determined by one day of observation in each class, as described in section II.

[c] The observed engagement rate was calculated by taking the ratio for each student of engaged time in mathematics (direct observation) and allocated time in mathematics (observer logs), as described in section II.

[d] Standard deviations are shown in parentheses.

unadjusted teacher ratings and is attributable entirely to the adjustment factor.

The adjusted teacher ratings of student attentiveness and academic status are the same variables in both the reading and the mathematics analyses. Therefore, the correlations between adjusted teacher ratings and academic status were all positive and large for mathematics just as for reading. However, the observed engagement rates for mathematics showed even lower correlations with academic status than did the observed rates for reading. These correlations for mathematics were essentially zero. Hence, the mathematics data provide confirmation of the conclusion that teacher ratings of student attentiveness were more strongly related to academic status than to observed engagement rates, while observed engagement rates were unrelated to academic status.

These findings support the conclusion that teacher ratings of student engagement do not provide useful data for analyses of instructional time because these ratings are strongly influenced by teacher perceptions of student aptitude. Therefore, analyses relating instructional time to student learning outcomes did not use teacher ratings of student attentiveness. However, the observed engagement rates included in the data discussed above were used in analyses of instructional time and student learning to adjust teacher records of allocated instructional time. Brief discussion of some of these analyses is presented below.

Predictive Validity

It is not the purpose of this paper to analyze the relationship between instructional time and student learning. Therefore, complete data from regression analyses conducted with time and learning data will not be presented or discussed here. Nevertheless, notation of a few of the

findings of these regression analyses will be made here as a means of examining the relative validity of different sources of data on instructional time in terms of their predictive relationships to student learning.

The foregoing discussion indicated that teacher ratings of student engagement are probably more closely related to student aptitude than they are to student engagement itself. However, estimates of student engagement rates based on the ratio of observed engaged time to observed allocated time are potentially useful for the purpose of adjusting teacher records of allocated instructional time. These observed engagement rates were obtained over the shorter (two-week) OA-OB period of the study. The teacher records of allocated time for the longer (eight-week) A-B period of the study were adjusted using these observed engagement rates. The adjusted teacher records were used in regression analyses relating instructional time to student learning. Comparisons between analyses using adjusted teacher records of allocated time and analyses using unadjusted teacher records provide some indication of the utility of observed engagement rates as a means of improving teacher records of allocated time.

These analyses regressed student posttest scores on three variables: the student's pretest score, his entering academic status, and his instructional time in the related content area. The unique variance accounted for by the instructional time variable was calculated for each analysis. Comparisons of analyses using adjusted and unadjusted teacher records of allocated instructional time indicated that a greater percentage of unique variance is accounted for by records of instructional time that have been adjusted using observed engagement rates. Inconsistent results

were obtained for the different subtests in reading. However, adjusted records of instructional time accounted for more variance in total reading scores in both analyses conducted on these scores (one with all subjects pooled and one with subjects pooled within each class). The adjusted instructional time accounted for more variance in five out of the six regression analyses conducted with teacher records of instructional time in mathematics.

It should be noted that the utility of the observed engagement adjustments for analyses with subjects pooled within each class, in addition to analyses with all subjects pooled, provides some indication that the engagement adjustments accounted for individual differences between students within classes as well as differences between classes in terms of engagement rates.

These findings suggest that teacher records of instructional time are more strongly related to student learning when they are adjusted using observed student engagement rates. The primary advantage of using the adjusted records of allocated time, rather than simply observing engaged time directly, is that the adjusted records provide an estimate of engaged instructional time over longer periods of time (such as the eight-week A-B period), where the more expensive procedure of direct observation is required only for some shorter period during which engagement rates are observed (such as the two-week OA-OB period).

It would also be desirable to examine the relative validity of teacher records of instructional time (either adjusted or unadjusted) and observed engaged instructional time, in terms of their predictive relationships to student learning. Unfortunately, however, the data set considered here does not lend itself to this comparison. The teacher records of allocated

time were obtained over an eight-week period, whereas the observation of engaged time was obtained over only a two-week period. It was found that very little student learning occurred over the shorter two-week period. Therefore, there was little to be predicted with observed engaged time.

Despite this difficulty, it is possible to make some rather rough comparisons between teacher records of allocated time and observed engaged time for instructional time and student learning in one mathematics subtest, that for place value. This was the only subtest in either reading or mathematics where any student learniing was detected over the two-week period (except one reading subtest where the average student gain score was much less than one standard deviation). Observed engaged time in place value accounted for 10.8 percent of unique variance on student posttest scores in place value (subjects pooled). Adjusted teacher records of allocated time accounted for 6.4 percent, while unadjusted allocated time accounted for less than one percent of unique posttest variance. However, the results for the teacher records may be due to inaccurate record-keeping by teachers of their instruction related to place value. When several related mathematics content categories were combined, adjusted teacher records accounted for 16.9 percent of the posttest variance for place value, while the unadjusted records accounted for 7.8 percent of this variance. The same combination of categories for observed engaged time resulted in only 3.1 percent of unique variance accounted for by time. Therefore, these data support the conclusion that teacher records of instructional time are comparable to observed engaged time. This appears to be particularly true when the teacher records are adjusted with observed student engagement rates and when combinations of teacher record

categories are used to compensate for possible inaccuracies in these records.

It should be noted that the conclusions above are tentative, and should be treated with some degree of scepticism. The data discussed include only one fifteen item subtest. Furthermore, the periods of instruction compared (the eight-week A-B period and the two-week OA-OB period) are quite dissimilar in length. Therefore, these results are only preliminary suggestions of what might be expected of teacher records in relationship to direct observation. Nevertheless, the economical researcher may well be encouraged to use teacher records of allocated instructional time, adjusted using observed student engagement rates.

IV SUMMARY AND CONCLUSIONS

Alternative procedures have been presented for collecting instructional time data. These include the use of teachers and the use of observers as alternative sources of data, as well as allocated instructional time and engaged instructional time as alternative forms for the data. Alternative procedures for estimating student rates of engagement have also been presented. Instrumentation and procedures for using that instrumentation have been described in some detail.

Data have been presented and discussed as a means of examining the relative utility of these various procedures. Inter-observer reliability data were examined for the observation of engaged instructional time. It was found that engaged instructional time in reading and mathematics content areas can be observed at a level of reliability that is suitable for most research purposes.

Next, data were presented comparing teachers and observers as alternative data sources for the same instructional time. In general, it was found that data on allocated instructional time obtained from teachers shows reasonably high positive correlations with both allocated and engaged time obtained from observers. Estimates of student rates of engagement, obtained from teachers, were used to adjust teacher records of allocated time, thereby providing teacher records of engaged time. It was found that teacher records of engaged time in reading instruction were generally no more highly correlated with observational data on engaged time than were teacher records of allocated time. Moreover, teacher records of engaged time in mathematics instruction were generally lower in

correlation with observational data on engaged time than were teacher records of allocated time. Therefore, it may be concluded that teacher estimates of student engagement rates are not useful for the purpose of obtaining records of engaged instructional time from teachers. It should be noted that this conclusion treats the observational data on engaged time as the criterion for the validity of the data from teachers. The fact that extensive training was conducted with the observers and that acceptable inter-observer reliability was obtained does support the use of the observational data as the criterion for engaged instructional time. However, further analyses were also conducted in examination of the validity (or lack thereof) of the teacher estimates of student engagement rates.

Intercorrelations were determined for teacher estimates of engagement rates, observer estimates of engagement rates, and academic status, a general measure of student achievement in reading. It was found that the teacher estimates of engagement were more highly correlated with academic status than with observer estimates of engagement. The observer estimates of engagement were obtained separately for reading instruction and for mathematics instruction. The observer estimates for reading showed a low positive correlation with academic status, lower than that for the teacher estimates. The observer estimates for mathematics instruction showed essentially no correlation (zero correlation) with academic status.

These findings support the previous conclusion that the teacher estimates of student engagement rates were not as valid as those obtained by observation. The measure of academic status can be seen as a general

indication of student aptitude for school achievement. This would suggest that the teacher ratings of student engagement were strongly influenced by teacher perceptions of student aptitude or level of achievement. The comparatively low positive correlation between academic status and observer estimates of engagement in reading, however, can be explained as a function of the effect of engagement in reading upon achievement in reading. This interpretation is supported by the lack of correlation between academic status (reading achievement) and observer estimates of engagement in mathematics. One would not expect engagement in mathematics instruction to have an effect upon reading achievement. Hence, it appears that teachers have difficulty estimating student engagement rates independently of student aptitude for academic achievement. Observer estimates of engagement are more likely to be valid.

There was also some discussion of the relative validity of teacher and observer instructional time data in terms of predicting student learning outcomes. The available data do not warrant more than tentative conclusions. Nevertheless, there was some indication that teacher and observer sources of instructional time data are comparable in terms of their association with student learning outcomes. In addition, the teacher records of allocated instructional time seemed to be more highly associated with student learning when they were adjusted using observed engagement rates and when combinations of instructional time categories were used to compensate for possible miscategorization of time by teachers.

This body of data indicates that teacher records of allocated time provide a relatively economical source of instructional time data. Furthermore, these records of allocated time can be adjusted using observed estimates of student engagement rates, thereby providing data on engaged

instructional time. The observed engagement rates obtained here were based upon two weeks of observation. However, it is likely that fewer days of observation, sampled at different points throughout an academic term, would provide adequate estimates of engagement more economically.

It is hoped that these findings will provide some guidance toward optimal procedures for collecting instructional time data.

REFERENCES

Far West Laboratory for Educational Research and Development. Program plan for continuation of Phase III-A: Beginning Teacher Evaluation Study. San Francisco, 1975.

Filby, N. N., & Dishaw, M. Refinement of reading and mathematics tests through an analysis of reactivity (Technical Report III-6). Beginning Teacher Evaluation Study, Far West Laboratory for Educational Research and Development, San Francisco, 1976.

Flanders, N. Analyzing teaching behavior. Reading, Massachusetts: Addison-Wesley, 1970.

Gump, P. The classroom behavior setting: Its nature and relation to study behavior. U.S.O.E. Final Report. Lawrence, Kansas: University of Kansas, 1967. (mimeographed)

Stallings, J. A., & Kaskowitz, D. Follow-through classroom observation evaluation, 1972-1973. Menlo Park, California: Stanford Research Institute, 1974.

Wiley, D. E., & Harnischfeger, A. Explosion of a myth: Quality of schooling and exposure to instruction, major educational vehicles. Educational Researcher, 3(4), 1974, 7-12.

Appendix A

Reading Content Categories

and

Examples of Teacher Logs

Specific Content Categories for Grade 2 Reading Instruction

Specific Content Category Number	Specific Content Category Name	General Content Category Number	Observation Content Category Number
	Decoding		
1	Single consonants	2	2
2	Consonant blends and digraphs	2	2
3	Variant consonants (c,g)	2	2
4	Vowels - short	2	2
5	Vowels - final e pattern - long vowels	1	1
6	Vowels - digraphs	1	1
7	Vowels - dipthongs	2	2
8	Vowels - vowels + r (car)	2	2
9	Complex, multi-syllabic	2	2
10	Silent letters	2	2
11	Sound substitution tasks	2	2
58	Spelling	2	2
14	Other decoding	2	2
	Context Clues		
15	Choosing word(s)which fit gram. context	3	5
16	Choosing word(s)which make best sense (semantic appropriateness)	3	5
17	Choosing correct form of word	3	5
18	Choosing word with correct initial cons.	3	5
19	Choosing correct pronoun	3	5
20	Other context clues	3	5
	Word Structure		
21	Compound words	4	3
22	Identification of root words	5	4
23	Prefixes - meaning and use	5	4
24	Suffixes - meaning and use	5	4
25	Contractions	5	4
26	Syllables	5	4
27	Other word structure	5	4
	Word Meaning		
28	Synonyms	6	5
29	Antonyms	6	5
30	Vocabulary building	6	5
31	Pronoun reference	6	5
32	Multi-meaning words in context	6	5
33	Unfamiliar words in context	6	5
34	Figurative language	6	5
35	Other word meaning	6	5

Comprehension

36	Understanding event detail	7	5
37	Understanding description	7	5
38	Understanding relationships	7	5
39	Understanding main idea	7	5
40	Literal recall	7	5
41	Translation of ideas	7	5
42	Synthesis of ideas, inference	7	5
43	Going beyond the text, prediction	7	5
44	Recognizing facts and opinions	7	5
45	General comprehension	7	5
46	Understanding directions	7	5
47	Picture interpretation to aid comprehension	7	5
51	Understanding signs	7	5
52	Understanding letters	7	5

Areas Related to Reading

48	Dictionary skills		
49	Reference sources in books (table of contents, index, glossary)	8	7
50	Choosing reference sources (dictionary, encyclopedia, card catalog)	8	7
53	Understanding Maps	8	7
54	Understanding Graphs	8	7
59	Grammar	8	7
60	Creative writing	8	7

Reading Practice

12	Sight words	9	6
13	Automaticity of word recognition	9	6
55	Reading for different purposes	9	6
56	Oral reading	9	6
57	Reading for enjoyment	9	6
61	Reading in content areas	9	6
62	Silent reading	9	6
67	Music (reading lyrics)	9	6

Miscellaneous

63	Listening (to story or tapes)	10	-
64	Penmanship and copying	10	-
65	Standardized tests	10	-
66	Foreign language	10	-
68	Dramatics (plays, choral reading...)	10	-

General Content Categories for Grade 2 Reading Instruction

General Content Category Number	General Content Category Name	Observation Content Category Number
1	Long vowels	1 (RL)
2	Other decoding	2 (RD)
3	Context clues	5 (RM)[a]
4	Compound words	3 (RC)
5	Other word structure	4 (RS)
6	Word meaning	5 (RM)[a]
7	Comprehension	5 (RM)[a]
8	Areas related to reading	7 (RO)
9	Reading practice	6 (RP)
10	Miscellaneous	--

[a] Observation content category 5 included general content categories 3, 6 and 7.

READING GLOSSARY

I. DECODING (Knowledge and use of letter-sound correspondence)

SCC 1 Single consonants -
Sounds of single consonants in any position in a word.
Examples: b, c, d, . . .

SCC 2 Consonant blends and digraphs
Blends include st, bl, tr, . . .
Digraphs include ch, sh, th, wh.

SCC 3 Variant consonants
A comparison of several sounds possible for a single consonant.
Examples: "c" in cat vs city, "g" in goat vs giant

SCC 4 Vowel - short
Regular short sound of a, e, i, o, and u

SCC 5 Vowel - final e pattern
Long vowel sound when word ends with e, as in rope

SCC 6 Vowel digraphs
Include ee, ea, ai, oa, and ay

SCC 7 Vowel dipthongs
Include oi, oo, ou, oy, au, and aw

SCC 8 Vowel plus r
Vowel sound modified by following consonant r
Examples: ar, er, ir, or, ur, air, ear

SCC 9 Complex, multi-syllabic
Decoding of multi-syllabic words, includes internal patterns, syllable influence on vowel decoding

SCC 10 Silent consonants
Letters which are not sounded in a word
Examples: comb, knit

SCC 11 Sound substitution tasks
Substituting one sound for another to create a new word.
Example: fan, _an, p, pan

SCC 12 Sight words
Recognition of common words, especially function words (the, of, to, would, could, were) and words with irregular spelling (are, come, put)

SCC 13 Automaticity of word recognition
Practice to improve speed of word recognition, so that the process becomes automatic.

II. CONTEXT CLUES

Context clues involve using the context of a phrase, sentence, or story to help identify a word or to predict a missing part. Different types

of context clues emphasize different aspects of the linguistic context or of the word to be identified.

SCC 15 Choosing word(s) which fit the grammatical context.
Father is sleeping _____ the bed.
night
in
warm

SCC 16 Choosing the word(s) which make best sense in the blank.
The _____ lives in the royal palace with her father.
princess
prince
sister

SCC 17 Choosing the correct form of a word.
Both of the _____ are asleep.
baby
babying
babies

SCC 18 Choosing the word with the correct initial consonant.
Don't _____ the milk.
sill
spill
still

SCC 19 Choosing correct pronoun.
John dropped his book and then picked _____ up.
them
it
him

III. WORD STRUCTURE

SCC 21 Compounds
Words formed by combining two smaller words - "mailbox"

SCC 22 Identification of root words
Recognizing the root word in a derived form - "playing" root = play

SCC 23 Prefixes -
Include re-, un-, dis-, pre-, . . .

SCC 24 Suffixes
Include grammatical endings like -s, -ed, and -ing and other suffixes like -ly, -ful, -ness, -less . . .

SCC 25 Contractions
do not - don't

SCC 26 Syllables - separation of a word into sound units preamble - pre am ble

IV. WORD MEANING

SCC 28 Identifying words with similar meanings - quick = fast

SCC 29 Antonyms
Identifying words with opposite meanings - large vs. small

SCC 30 Vocabulary building
Learning word meanings

SCC 31 Pronoun reference
Identifying the referent of a pronoun.
"John washed his car." his=John's

SCC 32 Multi-meaning words in context
Identifying the specific meaning of a word in a particular context.
I cut my hand on a piece of paper.
a. part of a clock
b. part of a person
c. give something

SCC 33 Unfamiliar words in context
Deducing the meaning of an unfamiliar word through its use in context.
The car was so badly entrenched in the mud that we had to call a tow truck.

a. stuck
b. built
c. dirty

SCC 34 Figurative language
Recognizing the meaning of a word or phrase used in a nonliteral sense, including simile, metaphor, and idiomatic expressions.
The soldier fought like a tiger to protect his home.
a. in a striped uniform
b. with sharp claws
c. bravely and fiercely

Oh, how Peter wished he could whistle! Peter saw his friend Sam playing with a dog. Whenever Sam whistled, the dog ran straight to him. Peter wished he could do that trick with his own dog, Willie. Peter tried and tried to whistle, but he just couldn't.

Peter went into his house and put on his father's old hat, to make himself feel more grown-up. He looked into the mirror to practice whistling. Still no whistle!

The next day Peter went outside to play. He sat on the front steps and tried to whistle. Then Peter saw his dog coming. Quick as a wink, Peter hid behind the stairs. He wanted to surprise Willie with a whistle. Peter puffed up his cheeks. He blew and blew and blew. Suddenly, out came a real whistle. Willie stopped and looked around to see who was making the noise.

"It's me," Peter shouted. He jumped out from behind the stairs. Willie raced straight up to him.

The following illustrations refer to the story above.

V. COMPREHENSION

SCC 36 Understanding event detail. What did Peter put on?

SCC 37 Understanding description - Where did Peter hide?
How did Peter feel at the end of the story?

SCC 38 Understanding relationships - What happened first?
Why did Willie stop and look around?

SCC 39 Understanding the main idea - What is the story mostly about?
What lesson can we learn from the story?

SCC 40 Literal recall - recall of information exactly as stated in the story.
What did Peter wish he could do?
a. have a dog
b. whistle
c. go to school

SCC 41 Translation of ideas
Recognizing ideas stated in different words; ability to paraphrase; recall of information when ideas are restated.

What happened when Sam whistled?
a. Peter went over to see Sam
b. A dog went over to see Sam
c. Peter whistled too

SCC 42 Synthesis of ideas, inference
Ability to integrate information from different points in a text; understanding ideas directly implied by a text.
What trick did Peter want to do with his dog?
a. teach Willie to whistle
b. put an old hat on Willie
c. whistle to call Willie

SCC 43 Going beyond the text, prediction
Relating the text to one's own knowledge and experience; supplying from experience information not directly given in a text. Includes predicting what might come next in a story.
How did Peter feel when Willie came running?
a. happy
b. scared
c. mad

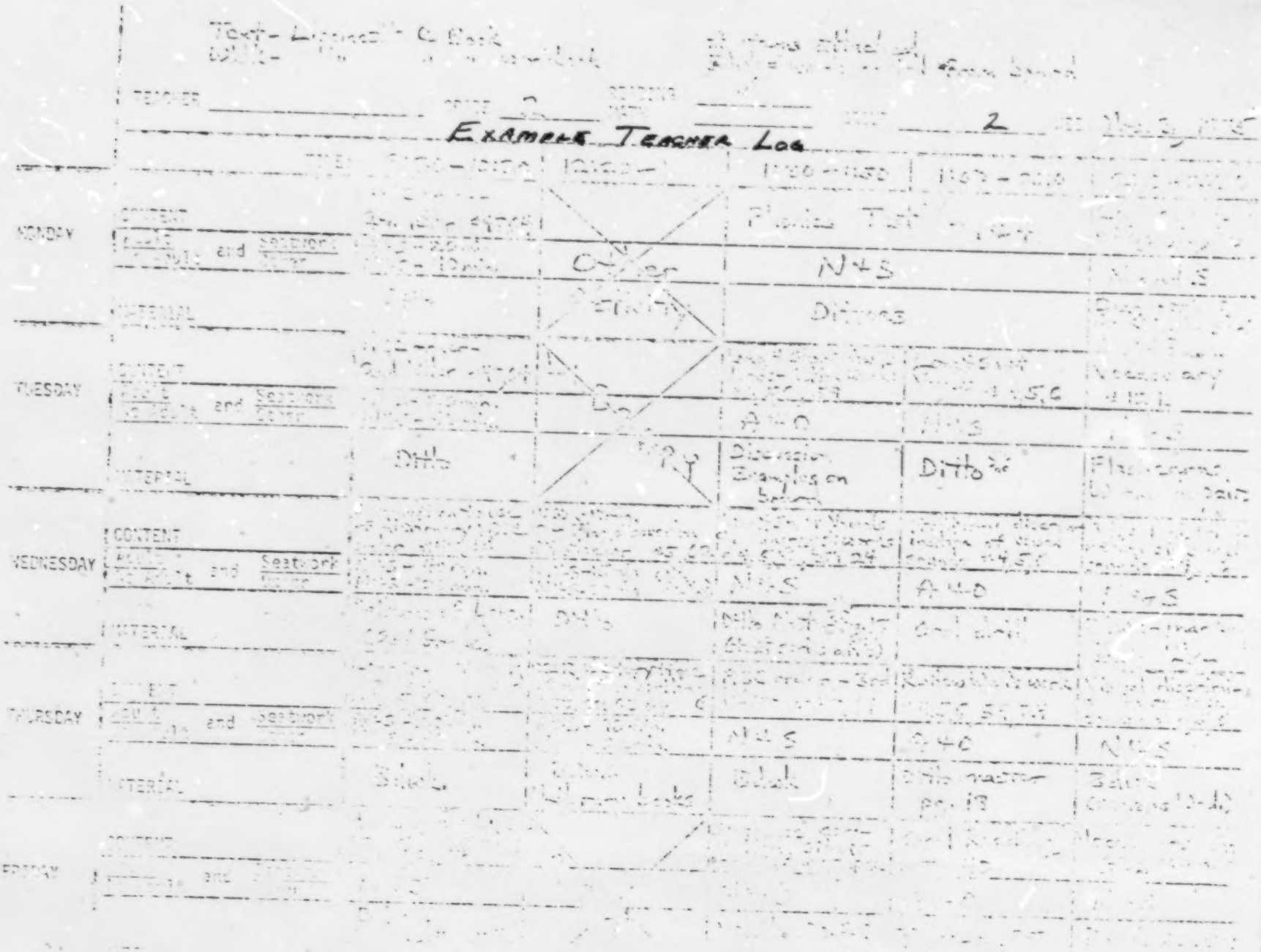
SCC 44 Recognizing facts and opinions
Evaluating statements and the basis for their acceptance. Included evaluating the qualifications of a speaker.
Which of the following is a <u>fact</u> rather than an opinion?
a. The Etruscans built cities long ago.
b. The jewelry made by the Etruscans was the most beautiful ever made.
c. Historians do not know as much as archeologists do.

SCC 45 General Comprehension
Silent reading or general reading practice, where comprehension involves a mixture of the facets above: (Please use one or more of the specific categories, if possible.)

Example: Attendance/Group Composition Record

(READING) MATH (circle one) Teacher No. 3 Grade 2

Student's Name	Group	Week of October 27-31, 1975				
		M	T	W	Th	F
1. ID # 239	3					
2. ID # 240	2					
3. ID # 241	2	Absent				
4. ID # 242	3					
5. ID # 243	3			Absent		
6. ID # 247	2					
7. ID # 251	3				No Absences	No Absences
8. ID # 252	2					
9. ID # 253	2					
10. ID # 254	3					
11. ID # 255	3					
12. ID # 256	2					
13. ID # 257	2			Absent		
14. ID # 258	2					
15. ID # 259	2					
16. ID # 260	3					
17. ID # 262	3					
18. ID # 263	2					
19. ID # 264	2					
20. ID # 265	3					
21. ID # 266	3					
22.						
23.						
24.						
25.						
26.						
27.						
28.						
29.						
30.						
31.						
32.						
33.						
34.						


EXAMPLE TEACHER LOG
MONDAY
TUESDAY
WEDNESDAY
THURSDAY
FRIDAY
Ditto
N.S
Ditto
N.S
N.S

Appendix B

Mathematics Content Categories

and

Examples of Teacher Logs

General and Specific Content Categories for
Grade 2 Mathematics Instruction

Specific Content Category Number	Category Name	General Content Category Number
	Computation	
1	Addition without regrouping	1
2	Addition with regrouping	2
3	Subtraction without regrouping	3
4	Subtraction with regrouping	4
5	Multiplication - with both factors being less than 10	9
6	Speed tests/timed drill in addition	1
7	Speed tests/timed drill in subtraction	3
8	Number sentences involving equalities and inequalities	6
9	Family of facts/renaming numerals equation form	1,3*
10	Number patterns/sequences	6
25	Missing addends - both in addition and subtraction	1,3*
11	Other - computation**	10
	Concepts	
12	Numerals and ordinals	6
13	Place value with compact or expanded notation	5
14	Fractions involving sets, regions, or lines (1/4,1/3,1/2,2/3,3/4)	9
15	Properties (associative, commutative, and identity elements)	6
16	Associative property with expanded notation	5,6*
17	Money	9
18	Linear measurements	7
19	Measurement concepts: order, capacity, conservation of length	7
20	Geometric figures:	8
21	Curves and points	8
26	Developmental activities	6
22	Other - concepts**	10

Specific Content Category Number	Category Name	General Content Category Number
	Applications	
23	Word problems	9
27	Standardized tests	10
24	Other - applications**	10

*Specific content categories 9,16, and 25 are logically related to two general content categories. In each case time in a specific content category was divided equally and assigned to the appropriate general content categories.

**Time in specific categories 11,22,24 was assigned to general content category 10 if it was not clear that the event could be assigned to general content categories 1-9.

General Content Categories

General Content Category Number	Category Name	Observation Content Category Number
1	Addition without regrouping	1
2	Addition with regrouping	2
3	Subtraction without regrouping	3
4	Subtraction with regrouping	4
5	Place value	5
6	Number system	6
7	Measurement	
8	Geometry	
9	Word problems	
10	Other	

Glossary: Selected Specific Content Categories

SCC 8 Number sentences: equalities and inequalities
Determining what sign is missing in an equation or number sentence.
Primarily involving the signs of $>$, $<$, and $=$. But may also involve + and −.

Examples: $5 \bigcirc 4$
$2 + 1 \bigcirc 3$
$3 \bigcirc 4 = 7$

SCC 9 Family of Facts (renaming numberals, equation form)

Example: Given 2 addends and a sum, write all the equations possible.

$3 + 2 = 5$ $\quad$ $5 - 2 = 3$
$2 + 3 = 5$ $\quad$ $5 - 3 = 2$

Write n equations renaming the numeral 7.

$4 + 3 = 7$ $\quad$ $9 - 2 = 7$
$5 + 2 = 7$ $\quad$ $8 - 1 = 7$ $\quad$ etc.
$6 + 1 = 7$ $\quad$ $10 - 3 = 7$

SCC 10 Number patterns
Emphasis is on completing the pattern and/or discovering the rule.
May ask "What is the next number?" or "What rule did you use to find the missing number?"

Examples: a. 2, 4, 6, ?, ?, 12 (series)

b.

2	3	4	?	?
4	6	?	?	12

(input/output function)

c.

4	?	10
3	5	?
?	11	18

(magic squares)

d. [(1,3) (2,4) (?,5) (?,?)] (set of number pairs)

SCC 15 Properties
Situations often arise in the teaching of basic facts, computation, etc. where particular items may illustrate certain basic properties. However, for the log, we are interested in lesson segments where attention is given or drawn to a particular property or the lesson involves a series of computational items which involve application/use of a property.

SCC 15 (cont'd)

1. Commutative (order): emphasis or attention at this stage is focused on the reversability of order.

   Examples: $2 + 4 = 4 + 2$
   $3 \times 2 = 2 \times 3$
   or $6 + 1 = 7,\ 1 + 6 = 7$ (When these equations appear together and attention is drawn to reversability)

2. Associative (grouping): manner in which numbers are grouped does not affect sum.

   Examples: $(2 + 3) + 5 = 5 + 5$
   $2 + (3 + 5) = (2 + 3) + 5$

   or can be combined with expanded notation:

   $78 + 2 = 70 + (8 + 2)$
   $= 70 + 10$
   $= 80$

   $38 = 30 + 8$ so
   $38 - 5 = (30 + 8) - 5$
   $= 30 + (8 - 5)$
   $= 30 + 3$
   $= 33$

3. Inverse: Addition is inverse of subtraction.

   Examples: $13 + 1 = 14$
   $14 - 1 = 13$

   $\begin{array}{r} 91 \\ -47 \\ \hline 44 \end{array}$ $\begin{array}{r} 44 \\ +47 \\ \hline 91 \end{array}$

4. Identity elements: zero for addition and subtraction, one for multiplication.

   Examples: $3 + 0 = 3$
   $4 - 0 = 4$
   $1 \times 6 = 6$

Example: Attendance/Group Composition Record

READING (MATH) (circle one) Teacher No. 3 Grade 2

Student's Name	Group	Week of October 27-31, 1975				
		M	T	W	Th	F
1. ID # 239	3					
2. ID # 240	2					
3. ID # 241	2	Absent				
4. ID # 242	3					
5. ID # 243	3			Absent		
6. ID # 247	2					
7. ID # 251	3					
8. ID # 252	2				No Absences	No Absences
9. ID # 253	2					
10. ID # 254	3					
11. ID # 255	3					
12. ID # 256	2					
13. ID # 257	2			Absent		
14. ID # 258	2					
15. ID # 259	2					
16. ID # 260	3					
17. ID # 262	3					
18. ID # 263	2					
19. ID # 264	2					
20. ID # 265	3					
21. ID # 266	3					
22.						
23.						
24.						
25.						
26.						
27.						
28.						
29.						
30.						
31.						
32.						
33.						
34.						

EXAMPLE: Teacher Log

TEACHER No. 3 GRADE 2 READING / MATH x GROUP 2 and 3 WEEK 10-27 to 31, 19

TIME: 10:50 - 11:50

MONDAY	CONTENT	#10 20 mins.	#13 20 mins.	#10 10 mins.
	Adult / No Adult and Seatwork / Other	A & S	A & S	A & O
	MATERIAL	Ditto write by 2's, 5's, 10's	Mod. math page 97-98	Board explanation by teacher with students participating at board
TUESDAY	CONTENT	#10	#1 #3	
	Adult / No Adult and Seatwork / Other	A & S	A & S	
	MATERIAL	Mod. math pages 99-100	Mod. math 101-102	
WEDNESDAY	CONTENT	#12	#10 20 mins.	#1 #3 15 mins.
	Adult / No Adult and Seatwork / Other	A & S	A & S	N & S
	MATERIAL	Ditto #2 Number names	Ditto: write to 100. Label even numbers with blue crayon	Color by number ditto(attached) + and - facts
THURSDAY	CONTENT	#12 #10	#10 #13	
	Adult / No Adult and Seatwork / Other	A S	A & S & O	
	MATERIAL	Ditto #3 used sticks sets and straws to do their work.	Correcting errors on papers - students working together in small groups with counters and sets of 10.	
FRIDAY	CONTENT			
	Adult / No Adult and Seatwork / Other	HALLOWEEN PARADE		
	MATERIAL			

APPENDIX C

Teacher Rating Forms

Rating of Student Attentiveness

The lesson plan logs tell us how much time was allocated by the teacher to different settings and objectives in reading or mathematics. But there is often a difference between the amount of time taken up by a lesson and the amount of time which is active learning time for a student. We would like to get a more accurate estimate of the amount of actual learning time for an individual student. This will be used as a "correction factor" in the interpretation of the log information.

A student who is paying attention	works actively on assignments participates or listens attentively during class discussion
A student who is not paying attention	talks to his neighbor daydreams draws pictures on his paper falls asleep walks around the room waits for help

Children differ in the amount of time they pay attention in class. Based on your observations of the children so far this year, please rate each child in your class as to the percentage of time that child generally pays attention. Think only of the subject matter for which you are keeping logs. If you are keeping math logs, think about how much of the time a student pays attention during math lessons. If you are keeping reading logs, think of the time you record as reading or reading related.

We suspect that children may differ in attentiveness depending on whether or not an adult is present. For this reason, we would like you to rate each child twice. First rate the children for settings in which an adult is present (either seatwork with an adult supervising or group work with an adult). Then rate the children a second time for settings with no adult present.

One way to go about this task is to go through the following steps:

1. Think of a typical 40 or 50 minute lesson period. Think first of settings where an adult is present to supervise and encourage attention. During what percentage of the time would a student by likely to pay attention to the lesson? On the form labeled "Adult Present," assign each student a rating.
2. Shift your thinking to a 40 or 50 minute period where the students are left to work on their own without an adult. What percentage of the time would a student be likely to pay attention under these conditions? Record your ratings on the form labeled "No Adult."

Use as many or as few of the categories as you wish to indicate the differences in attentiveness among your students. The descriptions below may help as guidelines.

91-100%	The child almost always attends to the learning task.
71-80%	The child sometimes loses time through temporary inattention or general classroom disruption but he tends to work more often than not.
51-60%	The child is as likely to be distracted as he is to work. Only about half the period is spent attending to the task.
21-30%	The child is frequently distracted and inattentive. Large periods of time may be lost through inattention. The child may be noticeably disruptive in class or may simply daydream a lot.
0-10%	The child almost never attends to the learning task.

Teacher's Name_____

Rating of Student Attentiveness

PUT A CHECK IN THE APPROPRIATE BOX

Adult Present

Student	0-10%	11-20%	21-30%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%

Teacher's Name_____

Rating of Student Attentiveness

PUT A CHECK IN THE APPROPRIATE BOX

Student	No Adult 0-10%	11-20%	21-30%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%